

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 1, 2004, 11:04:50 ; Search time 126 Seconds  
(without alignments)  
1080.856 Million cell updates/sec

Title: US-10-803-278-4  
Perfect score: 2501  
Sequence: 1 MLKFEAAKCVSGSTAISTY.....LGYHGDCLNLSLDEYWKNEK 482

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 1586107 seqs, 282547505 residues

Total number of hits satisfying chosen parameters: 1586107

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : A\_Geneseq\_28Jan04:\*  
1: Genesep1980s:\*  
2: Genesep1990s:\*  
3: Genesep2000s:\*  
4: Genesep2001s:\*  
5: Genesep2002s:\*  
6: Genesep2003as:\*  
7: Genesep2003bs:\*  
8: Genesep2004s:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	2408	96.3	470	7	ADC51184	ADC51184 Human cel
2	2408	96.3	645	4	AAU03546	Aau03546 Human pro
3	2408	96.3	645	5	AAE19147	Aae19147 Human kin
4	2408	96.3	645	7	ADC51186	ADC51186 Human cel
5	2405	96.2	645	5	ABG31081	Abg31081 Human nov
6	2182	87.2	609	7	ADC98091	Adc98091 Human KPP
7	1799.5	72.0	540	7	ADC98090	Adc98090 Human KPP
8	1652	66.1	497	5	ABP68951	Abp68951 Human pol
9	1491	59.6	467	5	ABP72299	Abp72299 Rat prote
10	916	36.6	357	7	ADC51182	ADC51182 Human cel
11	615	24.6	841	7	ADB37605	Adb37605 Neural th
12	613	24.5	841	7	ADB37570	Adb37570 Neural th
13	598.5	23.9	489	6	ABU54632	Abu54632 Human NOV
14	598	23.9	774	6	AAO26613	Aao26613 Serine/th
15	598	23.9	774	6	ABP71711	Abp71711 NEK-like
16	598	23.9	774	6	AAO16441	Aao16441 Human ser
17	598	23.9	774	6	ABP97691	Abp97691 Amino aci
18	593.5	23.7	425	6	ABU54633	Abu54633 Human NOV
19	592	23.7	506	6	ABU54631	Abu54631 Human NOV
20	590.5	23.6	489	7	ADC98067	Adc98067 Human KPP
21	587.5	23.5	507	6	ABP97688	Abp97688 Amino aci
22	585.5	23.4	654	5	AAU7929	Aau7929 Human pro
23	585	23.4	506	4	AAU78344	Aau78344 Human ser
24	585	23.4	506	5	ABP60668	Abp60668 Human ser
25	585	23.4	546	6	ABP96072	Abp96072 Human pro

26	584.5	23.4	511	6	ABP71712	Abp71712 NEK-like
27	584.5	23.4	511	6	ABP97690	Abp97690 Amino aci
28	584	23.4	506	5	AAE24136	Aae24136 Human kin
29	584	23.4	640	6	ABG72000	Abg72000 Human ser
30	584	23.4	683	5	AAU7928	Aau7928 Amino aci
31	582.5	23.3	1214	4	AAU39211	Aau39211 Human pol
32	582.5	23.3	1214	4	AAU07102	Aau07102 Human nov
33	582.5	23.3	1214	6	ABU08113	Abu08113 Human kin
34	582.5	23.3	1242	4	AAU39210	Aau39210 Human pol
35	582.5	23.3	1242	6	AAO16440	Aao16440 Human ser
36	581.5	23.3	649	4	AAU03545	Aau03545 Human pro
37	580	23.2	527	4	AAU79328	Aau79328 Human pro
38	580	23.2	527	5	ABP97224	Abp97224 Novel hum
39	575.5	23.0	403	6	ABU54635	Abu54635 Human NOV
40	572.5	22.9	631	6	ABU71710	Abu71710 Human NOV
41	563.5	22.5	399	6	ABU54634	Abu54634 Human NOV
42	563	22.5	112	7	ADC51180	Adc51180 Human cel
43	560	22.4	510	3	AAU68778	Aau68778 Amino aci
44	539	21.6	459	6	ABP71713	Abp71713 NEK-like
45	539	21.6	459	6	ABP97689	Abp97689 Amino aci

## ALIGNMENTS

## RESULT 1

ADC51184  
ID ADC51184 standard; protein; 470 AA.  
XX  
AC ADC51184;  
XX  
DT 18-DEC-2003 (first entry)  
XX  
DE Human cell-cycle related protein, SEQ ID 10.  
XX  
KW Human; cytostatic; cell-cycle related protein; nuclear export;  
KW nuclear-cytoplasm transport; cytotoxic; cell-cycle control;  
KW immunological disease; neurological disease; cancer.  
XX  
OS Homo sapiens.  
XX  
FN JP2003144168-A.  
XX  
PD 20-MAY-2003.  
XX  
PF 14-NOV-2001; 2001JP-00349158.  
XX  
PR 14-NOV-2001; 2001JP-00349158.  
XX  
PA (KOKU-) KOKURITSU YOBO EISEI KENKYUSHO.  
PA (KAGA-) KAGAKU GIJUTSU SHINKO JIGYODAN.  
DR  
DR WPI; 2003-818166/77.  
DR N-PSDB; ADC51183.  
PT Novel DNA or RNA coding a cell-cycle related protein which has nuclear  
PT export function, useful for screening substance that prevent or treat  
PT cell cycle abnormality diseases e.g. immunological disease.  
PS Disclosure; SEQ ID NO 10; 41pp; Japanese.  
XX  
XX The present invention relates to novel cell-cycle related protein such as  
CC NIMA (Never-In Mitosis, gene A)-related protein kinase of Nek 9 and  
CC coding sequences such as a cell-cycle related protein (ADC51176) having  
CC nuclear export function; cell-cycle related protein (ADC51178) having  
CC nuclear-cytoplasm transport function; cell-cycle related protein  
CC (ADC51180) having nuclear export function; cytotoxic function and  
CC transfer function in the nucleus; and/or cell-cycle related protein  
CC (ADC51182) having transfer function and cytotoxic function in nucleus,  
CC nuclear export function and nuclear cytoplasm transport function. The  
CC coding sequences for these proteins are given in ADC51175, ADC51177,  
CC ADC51179 and ADC51181. The sequences of the invention are useful for  
CC screening a substance which promotes or suppresses the transfer function

CC or cytotoxic function in the nucleus, nuclear-cytoplasm transport  
CC function and/or a cell-cycle control function. The sequences are also  
CC useful for treating or diagnosing an immunological disease, neurological  
CC disease or cancer. The present sequence was used to illustrate the  
CC invention.  
XX  
SQ Sequence 470 AA;

Query Match	96.3%; Score 2408; DB 7; Length 470;
Best Local Similarity	100.0%; Pred. No. 4.1e-177;
Matches 466; Conservative 0; Mismatches 0; Indels 0; Gaps 0;	
QY 1 MLKFOEAAKCVSGSTAIPTPKTLIARRVYVLOQKLGSGFGTVLVSDKKAKRGELKVL 60	
DB 1 MLKFOEAAKCVSGSTAIPTPKTLIARRVYVLOQKLGSGFGTVLVSDKKAKRGELKVL 60	
QY 61 KEISVGLNPNETVQANLEAQLLSKLDHPAIVKHFASFVEQDNFCIIITEYCEGRDLDDKI 120	
DB 61 KEISVGLNPNETVQANLEAQLLSKLDHPAIVKHFASFVEQDNFCIIITEYCEGRDLDDKI 120	
QY 121 QEYKQAGKIPFENQIIEFWFIQLLGVDMHERRILHRDLKSNVFLKNNLLKIGDFGVS 180	
DB 121 QEYKQAGKIPFENQIIEFWFIQLLGVDMHERRILHRDLKSNVFLKNNLLKIGDFGVS 180	
QY 181 LLMGSCDLATTLTGTPHYMSPEALKHQGYDTKSDIWSLACILYEMCMNHAFAGSNFLSI 240	
DB 181 LLMGSCDLATTLTGTPHYMSPEALKHQGYDTKSDIWSLACILYEMCMNHAFAGSNFLSI 240	
QY 241 VLKIVEGDTSPSLPERYPKELNAMESMLNKNPSLRPSAIEILKIPYLDEQLNLCRYSE 300	
DB 241 VLKIVEGDTSPSLPERYPKELNAMESMLNKNPSLRPSAIEILKIPYLDEQLNLCRYSE 300	
QY 301 MLEDKNLDCKQEAHIINAMQRIHLQTLRALSEVQKMTPRERMLRKLQAADKARKL 360	
DB 301 MLEDKNLDCKQEAHIINAMQRIHLQTLRALSEVQKMTPRERMLRKLQAADKARKL 360	
QY 361 KKIIVEKYENSKRMQELRSRNFQQLSDVVLHFKTHLKGMEKEEQEPGRSLSCSPQDE 420	
DB 361 KKIIVEKYENSKRMQELRSRNFQQLSDVVLHFKTHLKGMEKEEQEPGRSLSCSPQDE 420	

RESULT 2  
AAU03546  
ID AAU03546 standard; protein; 645 AA.  
AC AAU03546;  
XX  
XX 12-SEP-2001 (first entry)  
DE Human protein kinase #46.  
XX  
XX Human; protein kinase; PTK; STK; cancer; cardiovascular disease;  
KW metabolic disorder; immune related disease; neurological disorder;  
KW neurodegenerative disorder; inflammatory disorder; infectious disease;  
KW reproductive disorder.  
XX  
XX Homo sapiens.  
OS  
XX WO200138503-A2.  
PN  
XX 31-MAY-2001.  
XX  
XX 22-NOV-2000; 2000WO-US032085.  
PF  
XX 24-NOV-1999; 99US-0167482P.  
PR  
XX (SUGB-) SUGEN INC.  
XX  
XX Plowman GD, Whyte D, Manning G, Sudarsanam S, Martinez R;  
PI

PI Flanagan P, Clary D;  
XX  
DR WPI; 2001-343950/36.  
DR N-PSDB; AAS06746.  
XX  
PT Nucleic acids encoding human kinase polypeptides, useful for preventing  
PT diagnosing and/or treating e.g. cancer, immune, cardiovascular and  
PT neuronal-associated diseases, and microbial infections.  
XX  
PS Claim 7; Fig 2; 433pp; English.

XX AAU03501-AAU03557 represent novel human protein kinases #1-57. The novel  
CC protein kinases have been identified as members of the tyrosine or  
CC serine/threonine kinase (PTK and STK) families. The polynucleotides  
CC encoding protein kinases and the polypeptides may be used in the  
CC prevention, diagnosis and treatment of diseases associated with  
CC inappropriate kinase expression. For example, they may be used to treat  
CC cancers (especially cancers of haematopoietic origin), cardiovascular  
CC disease (e.g. atherosclerosis), metabolic disorders (e.g. diabetes),  
CC immune related diseases (e.g. rheumatoid arthritis), neurological  
CC disorders (e.g. schizophrenia), neurodegenerative disorders (e.g.  
CC Parkinson's disease), inflammatory disorders (e.g. asthma), infectious  
CC disease (e.g. HIV) and reproductive disorders (e.g. infertility).  
CC Additionally, polynucleotides encoding protein kinases may be used for  
CC gene therapy and as DNA probes in diagnostic assays. The protein kinase  
CC polypeptides may be used as antigens in the production of antibodies  
CC against the protein kinases and in assays to identify modulators of  
CC protein kinase expression and activity  
XX  
SQ Sequence 645 AA;

Query Match	96.3%; Score 2408; DB 4; Length 645;
Best Local Similarity	100.0%; Pred. No. 6.2e-177;
Matches 466; Conservative 0; Mismatches 0; Indels 0; Gaps 0;	
QY 1 MLKFOEAAKCVSGSTAIPTPKTLIARRVYVLOQKLGSGFGTVLVSDKKAKRGELKVL 60	
DB 1 MLKFOEAAKCVSGSTAIPTPKTLIARRVYVLOQKLGSGFGTVLVSDKKAKRGELKVL 60	
QY 61 KEISVGLNPNETVQANLEAQLLSKLDHPAIVKHFASFVEQDNFCIIITEYCEGRDLDDKI 120	
DB 61 KEISVGLNPNETVQANLEAQLLSKLDHPAIVKHFASFVEQDNFCIIITEYCEGRDLDDKI 120	
QY 121 QEYKQAGKIPFENQIIEFWFIQLLGVDMHERRILHRDLKSNVFLKNNLLKIGDFGVS 180	
DB 121 QEYKQAGKIPFENQIIEFWFIQLLGVDMHERRILHRDLKSNVFLKNNLLKIGDFGVS 180	
QY 181 LLMGSCDLATTLTGTPHYMSPEALKHQGYDTKSDIWSLACILYEMCMNHAFAGSNFLSI 240	
DB 181 LLMGSCDLATTLTGTPHYMSPEALKHQGYDTKSDIWSLACILYEMCMNHAFAGSNFLSI 240	
QY 241 VLKIVEGDTSPSLPERYPKELNAMESMLNKNPSLRPSAIEILKIPYLDEQLNLCRYSE 300	
DB 241 VLKIVEGDTSPSLPERYPKELNAMESMLNKNPSLRPSAIEILKIPYLDEQLNLCRYSE 300	
QY 301 MLEDKNLDCKQEAHIINAMQRIHLQTLRALSEVQKMTPRERMLRKLQAADKARKL 360	
DB 301 MLEDKNLDCKQEAHIINAMQRIHLQTLRALSEVQKMTPRERMLRKLQAADKARKL 360	
QY 361 KKIIVEKYENSKRMQELRSRNFQQLSDVVLHFKTHLKGMEKEEQEPGRSLSCSPQDE 420	
DB 361 KKIIVEKYENSKRMQELRSRNFQQLSDVVLHFKTHLKGMEKEEQEPGRSLSCSPQDE 420	
QY 421 ERWQGREESDEPTLENLPESQIPSPMDLHELESIVEDATSDLGYYH 466	
DB 421 ERWQGREESDEPTLENLPESQIPSPMDLHELESIVEDATSDLGYYH 466	

RESULT 3  
AAE19147  
ID AAE19147 standard; protein; 645 AA.  
XX  
XX AAE19147;

21-MAY-2002 (first entry)  
Human kinase polypeptide (PKIN-5) .  
Human; kinase polypeptide; PKIN-5; gene therapy; Addison's disease; leukaemia; immune disorder; lymphoma; melanoma; developmental disorder; acquired immunodeficiency syndrome; AIDS; allergy; anaemia; hypertension; ascirrhosis; Crohn's disease; rheumatoid arthritis; bursitis; atherosclerosis; cholelasis; hepatitis; psoriasis; Cushing's syndrome; cytostatic; cancer; lipid disorder; cardiovascular disorder; Niemann-Pick's disease; Niemann-Pick's disease; drug screening; transgenic animal; antiinflammatory; hepatotropic; hypotensive; anti-HIV; enzyme.  
Homo sapiens.  
Key Location/Qualifiers  
Domain 29..287  
FT FT /note="Eukaryotic protein kinase domain"  
Domain 108..121  
FT FT /note="Tyrosine kinase catalytic domain"  
FT FT 148..166  
FT FT /note="Tyrosine kinase catalytic domain"  
Domain 256..278  
FT FT /note="Tyrosine kinase catalytic domain"  
XX XX W0200208399-A2.  
XX XX  
XX XX 31-JAN-2002.  
XX XX  
XX XX 20-JUL-2001; 2001WO-US032092.  
XX XX  
XX XX 21-JUL-2000; 2000US-0220038P.  
XX XX 28-JUL-2000; 2000US-0222112P.  
XX XX 04-AUG-2000; 2000US-0225831P.  
XX XX 11-AUG-2000; 2000US-0224729P.  
XX XX  
XX XX (INCY-) INCYTE GENOMICS INC.  
XX XX (THOR/) THORNTON M.  
XX XX  
XX XX Thornton M, Yue H, Khan FA, Gururajan R, Hafalia AYA, Wallia NK, Patterson C, Ramkumar J, Gandhi AR, Policky JL, Baughn MR; Tribouley CM, Bandman O, Nguyen DB, Lu Y, Burford N, Lal P, Ding L, Yao MG, Elliott VS, Recipon SA, Kearney L, Lu DAM, Greenwald SR, Tang YT, Xu Y, Walsh RT, Gietzen KJ, Yang J, Hillman JL;  
XX XX WPI: 2002-206083/26.  
XX XX N-PSDB; AAD30552.  
XX XX  
XX XX New human kinase polypeptide, useful in diagnosis, prevention and treatment of cancer, immune disorder, growth and developmental disorder, cardiovascular disorder and lipid disorder.  
XX XX  
XX XX Claim 1; Page 138-140; 196pp; English.  
XX XX  
XX XX The present invention relates to an isolated human kinase polypeptide (PKIN) or a biologically active/immunogenic fragment of PKIN. PKIN is useful for diagnosing, treating and preventing cancer (e.g., leukaemia, lymphoma, melanoma), an immune disorder (e.g., acquired immunodeficiency syndrome (AIDS), Addison's disease, allergy, anaemia, asthma, Crohn's disease, rheumatoid arthritis), a growth and developmental disorder (e.g., bursitis, cirrhosis, hepatitis, psoriasis, Cushing's syndrome), a cardiovascular disorder (e.g., atherosclerosis, hypertension, myocardial infarction), and a lipid disorder (e.g., fatty liver, cholestasis, Gaucher's disease, Niemann-Pick's disease). PKIN is useful in a number of drug screening techniques and to analyse the proteome of a tissue or cell type. PKIN is useful for creating knockin humanised animals or transgenic animals to model human diseases, in somatic or germline gene therapy, to generate a transcript image of a tissue or cell type, for detecting differences in the chromosomal location due to translocation, inversion, etc., among normal, carrier or affected individuals, and as hybridisation probes for mapping naturally occurring genomic sequences. PKIN is useful

CC	in southern or northern analysis, dot blot or other membrane-based technologies, in PCR technologies, in dipstick, pin, multiformat enzyme linked immunosorbent (ELISA)-like assays and in microarrays utilising fluids or tissues from patients to detect altered PKIN expression. The present sequence is human PKIN-5
XX	Sequence 645 AA;
SQ	
Query Match	96.3%; Score 2408; DB 5; Length 645;
Best Local Similarity	100.0%; Pred. No. 6.2e-177;
Matches 466; Conservative	0; Mismatches 0; Indels 0; Gaps 0;
QY	1 MLKFOAAKCVSGSTAIPTPKTLIARRYVLQOKLGSGFGTVLYVSDDKAKRGEELKVL 60
Db	1 MLKFOAAKCVSGSTAIPTPKTLIARRYVLQOKLGSGFGTVLYVSDDKAKRGEELKVL 60
QY	61 KEISVGELPNPNETVOANLEAQLLSKLDPALVKFHFASFVEQDNFCIIITYCEGRDLDDKI 120
Db	61 KEISVGELPNPNETVOANLEAQLLSKLDPALVKFHFASFVEQDNFCIIITYCEGRDLDDKI 120
QY	121 QEYKQAGKIFPENQIITENFIQLLGVVDYMHERRIILHRDLKNVFLKNNLLIKIGDFGVS 180
Db	121 QEYKQAGKIFPENQIITENFIQLLGVVDYMHERRIILHRDLKNVFLKNNLLIKIGDFGVS 180
QY	181 LLMGSCDALTITLGTTHYMSPALKHOGYDTKSDIWSLACILIYEMCMNHAFAGSNFLSI 240
Db	181 LLMGSCDALTITLGTTHYMSPALKHOGYDTKSDIWSLACILIYEMCMNHAFAGSNFLSI 240
QY	241 VLKIVGDTPSLPERYPKELNAIMESMLNKPSPRLPSAIEILKIPYLDEQLQNLMCRYSE 300
Db	241 VLKIVGDTPSLPERYPKELNAIMESMLNKPSPRLPSAIEILKIPYLDEQLQNLMCRYSE 300
QY	301 MTLEDKNLDQCKEAAHI INAMOKRIHLOTLRALSEVQKMTPRERWLRKLQAADKARKL 360
Db	301 MTLEDKNLDQCKEAAHI INAMOKRIHLOTLRALSEVQKMTPRERWLRKLQAADKARKL 360
QY	361 KKIVEEKYEENSKRMQELRSRNFOQLSVDDLHEKTHLKGMEEKEQEPEGRLLSCSPODEDE 420
Db	361 KKIVEEKYEENSKRMQELRSRNFOQLSVDDLHEKTHLKGMEEKEQEPEGRLLSCSPODEDE 420
QY	421 ERWQGREESDBFTLENLPESQPIPSMDJHELESIVEDATSLGLGH 466
Db	421 ERWQGREESDBFTLENLPESQPIPSMDJHELESIVEDATSLGLGH 466
RESULT 4	
ADC51186	
ID	ADC51186 standard; protein; 645 AA.
XC	ADC51186;
AC	
DT	18-DEC-2003 (first entry)
XX	
XX	Human cell-cycle related protein, SEQ ID 12.
DE	
DE	
KW	Human; cytostatic; cell-cycle related protein; nuclear export;
KW	nuclear-cytoplasm transport; cytotoxic; cell-cycle control;
KW	immunological disease; neurological disease; cancer.
XX	
OS	Homo sapiens.
XP	
PN	JP2003144168-A.
XX	
PD	20-MAY-2003.
PF	
PF	14-NOV-2001; 2001JP-00349158.
XX	
PR	14-NOV-2001; 2001JP-00349158.
XX	
XX	(KOKU-) KOKURITSU YORO EISEI KENKYUSHO.
PA	(KAGA-) KAGAKU GIJUTSU SHINKO JIGYODAN.
XX	
DR	WPI: 2003-818166/77.

DR N-PSDB; ADC51185.  
 XX Novel DNA or RNA coding a cell-cycle related protein which has nuclear  
 PT export function, useful for screening substance that prevent or treat  
 PT cell cycle abnormality diseases e.g. immunological disease.  
 XX  
 XX Disclosure; SEQ ID NO 12; 41pp; Japanese.  
 XX  
 CC The present invention relates to novel cell-cycle related protein such as  
 CC NIMA (Never-in Mitosis, gene A)-related protein kinase of Nek 9 and  
 CC coding sequences such as a cell-cycle related protein (ADC51176) having  
 CC nuclear export function; cell-cycle related protein (ADC51178) having  
 CC nuclear-cytoplasm transport function; cell-cycle related protein  
 CC (ADC51180) having nuclear export function, cytotoxic function and  
 CC transfer function in the nucleus; and/or cell-cycle related protein  
 CC (ADC51182) having transfer function and cytotoxic function in nucleus,  
 CC nuclear export function and nuclear cytoplasm transport function. The  
 CC coding sequences for these proteins are given in ADC51175, ADC51177,  
 CC ADC51179 and ADC51181. The sequences of the invention are useful for  
 CC screening a substance which promotes or suppresses the transfer function  
 CC or cytotoxic function in the nucleus, nuclear-cytoplasm transport  
 CC function and/or a cell-cycle control function. The sequences are also  
 CC useful for treating or diagnosing an immunological disease, neurological  
 CC disease or cancer. The present sequence was used to illustrate the  
 CC invention.  
 XX  
 SQ Sequence 645 AA;

Query Match 96.3%; Score 2408; DB 7; Length 645;  
 Best Local Similarity 100.0%; Pred. No. 6.2e-177; Indels 0; Gaps 0;  
 Matches 466; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 MLKFEAAKCVSGSTAISTYPTKTLIARYVLOQKLGSGFGTVYLVSDKAKRGSELKVL 60  
 Db 1 MLKFEAAKCVSGSTAISTYPTKTLIARYVLOQKLGSGFGTVYLVSDKAKRGSELKVL 60  
 QY 61 KEISVGLNPNETVQANLEAQLSKLDHPAIVKPHASVEQDNFCIITEYCEGRDLDDKI 120  
 Db 61 KEISVGLNPNETVQANLEAQLSKLDHPAIVKPHASVEQDNFCIITEYCEGRDLDDKI 120  
 QY 121 QEYKQAGKIPFENQIIEWFTQLLGVDMYHERRILHRDLKSNVFLKNNLLKIGDFGVSR 180  
 Db 121 QEYKQAGKIPFENQIIEWFTQLLGVDMYHERRILHRDLKSNVFLKNNLLKIGDFGVSR 180  
 QY 181 LMGSCLATTLTGTHYNSPEALKHQGYDTKSDIWSLACILYEMCCMHAFAGSNFLSI 240  
 Db 181 LMGSCLATTLTGTHYNSPEALKHQGYDTKSDIWSLACILYEMCCMHAFAGSNFLSI 240  
 QY 241 VLKIVSGDTPSLPERYPKELNAMESMLKNPFLRPSAIEILKIPYLDQNLQMCRYSE 300  
 Db 241 VLKIVSGDTPSLPERYPKELNAMESMLKNPFLRPSAIEILKIPYLDQNLQMCRYSE 300  
 QY 301 MTLEDKXLDCKQEAHIINAMQKRIHLQTLRALSEVQKMTPRERMLRKLQAADKAKKL 360  
 Db 301 MTLEDKXLDCKQEAHIINAMQKRIHLQTLRALSEVQKMTPRERMLRKLQAADKAKKL 360  
 QY 361 KKIIVKYEENSKRMQELSRNFQQLSVDLVLEKTHLKGMBEKEQPEGRILSCSPQDEDE 420  
 Db 361 KKIIVKYEENSKRMQELSRNFQQLSVDLVLEKTHLKGMBEKEQPEGRILSCSPQDEDE 420  
 QY 421 ERWQGRREESDEPTLENLPESQIPSMDLHELESIVEDATSDLGX 466  
 Db 421 ERWQGRREESDEPTLENLPESQIPSMDLHELESIVEDATSDLGX 466

RESULT 5  
 ABG31081  
 ID ABG31081 standard; protein; 645 AA.  
 XX  
 XX ABG31081;  
 XX 21-OCT-2002 (first entry)  
 XX

DE Human novel serine/threonine serine kinase.  
 XX Human; enzyme; serine/threonine protein kinase; inflammation; cancer;  
 KW arteriosclerosis; psoriasis; SNF kinase; immunogen; chromosome 3.  
 XX  
 OS Homo sapiens.  
 XX  
 PN US2002082189-A1.  
 XX  
 PD 27-JUN-2002.  
 XX  
 PF 07-DEC-2000; 2000US-00731231.  
 XX  
 PR 07-DEC-2000; 2000US-00731231.  
 XX  
 PA (GUEG/) GUEGLER K.  
 PA (KSTC/) KETCHUM K A.  
 PA (IPRA/) DI FRANCESCO V.  
 PA (BEAS/) BEASLEY E M.  
 XX  
 PI Guegler K, Ketchum KA, Di Francesco V, Beasley EM;  
 XX WPI; 2002-598989/54.  
 XX N-PSDB; ABK89295, ABK89296.  
 XX  
 PT New isolated human kinase peptide for detecting a modulator of the  
 PT peptide's expression, activity or function, that can be used to treat  
 PT disorders or disease.  
 XX  
 PS Claim 1; Fig 2; 321pp; English.  
 XX  
 CC The invention relates to an isolated kinase peptide, comprising, a novel  
 CC human serine/threonine protein kinase sequence, its allelic variant or  
 CC orthologue, where the variant is encoded by a nucleic acid molecule that  
 CC hybridises under stringent conditions to the opposite strand of kinase  
 CC cDNA or gene, or a fragment comprising 10 contiguous amino acids. Also  
 CC included are an antibody that selectively binds the kinase, a gene chip  
 CC comprising a nucleic acid (or its complement) which encodes the kinase, a  
 CC transgenic non-human animal comprising the nucleic acid, a nucleic acid  
 CC vector comprising the nucleic acid and a host cell comprising the vector.  
 CC The kinases used to identify a modulator of the kinase or to identify  
 CC an agent that binds to the kinase, which can be used to treat a disease  
 CC or condition e.g. inflammation, cancer, arteriosclerosis and psoriasis.  
 CC The nucleic acid encoding the kinase is used to produce the kinase. A  
 CC detection agent is used to detect the kinase and an oligonucleotide is  
 CC used to detect a nucleic acid encoding the kinase in a sample. The kinase  
 CC can be used: (a) to raise antibodies against the kinase or to elicit  
 CC another immune response; (b) as a reagent (including a labeling reagent)  
 CC in assays to determine levels of a kinase (or its binding partner or  
 CC ligand) in biological fluids; and (c) as markers for tissues in which the  
 CC corresponding kinase is expressed. The kinase and antibodies against it  
 CC can be used in pharmacogenomic analysis. The kinase can be used to treat  
 CC a disorder characterised by an absence of, inappropriate, or unwanted  
 CC expression of the kinase. The kinase, homologous to SNF kinases (not  
 CC defined) is expressed by a gene located on human chromosome 3. The  
 CC present sequence represents the novel human kinase  
 XX  
 SQ Sequence 645 AA;

Query Match 96.2%; Score 2405; DB 5; Length 645;  
 Best Local Similarity 99.8%; Pred. No. 1.1e-176;  
 Matches 465; Conservative 1; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 MLKFEAAKCVSGSTAISTYPTKTLIARYVLOQKLGSGFGTVYLVSDKAKRGSELKVL 60  
 Db 1 MLKFEAAKCVSGSTAISTYPTKTLIARYVLOQKLGSGFGTVYLVSDKAKRGSELKVL 60  
 QY 61 KEISVGLNPNETVQANLEAQLSKLDHPAIVKPHASVEQDNFCIITEYCEGRDLDDKI 120  
 Db 61 KEISVGLNPNETVQANLEAQLSKLDHPAIVKPHASVEQDNFCIITEYCEGRDLDDKI 120  
 QY 121 QEYKQAGKIPFENQIIEWFTQLLGVDMYHERRILHRDLKSNVFLKNNLLKIGDFGVSR 180



Db 121 QYKQAGKIPFNQIIEWFIQLLLGVDMYHERRILHRDLKSKNVFLKNNLLKIGDFGVS 180  
Qy 181 LLMGSCDLATLTGTPTPHMSPEALKHQGYDTKSDIWSLACILYEMCCMHAFAGSNFLSI 240  
Db 181 LLMGSCDLATLTGTPTPHMSPEALKHQGYDTKSDIWSLACILYEMCCMHAFAGSNFLSI 240  
Qy 241 VLKIVEGDTPSLPERYKELNAMESMLNKNPSLRPSAIEILKIPYLDQQLNLMCRYS 300  
Db 241 VLKIVEGDTPSLPERYKELNAMESMLNKNPSLRPSAIEILKIPYLDQQLNLMCRYS 300  
Qy 301 MTLEDKNLDCQKEAAHIINAMQKRIHLQTLRALSEVQKTPRERMLRKLQAADEKARKL 360  
Db 301 MTLEDKNLDCQKEAAHIINAMQKRIHLQTLRALSEVQKTPRERMLRKLQAADEKARKL 360  
Qy 361 KKIIVEKYENSKEMLSRNFQQLSDVVLHEKTHLKGMEKEQPEGRSLSCSPODEDE 420  
Db 361 KKIIVEKYENSKEMLSRNFQQLSDVVLHEKTHLKGMEKEQPEGRSLSCSPODEDE 420  
Qy 421 ERWQGREESDEPTLENLPESQIPSMDLHELESIVEDATSDLGYN 466  
Db 421 ERWQGREESDEPTLENLPESQIPSMDLHELESIVEDATSDLGYN 466

RESULT 6  
ADC99091  
ID ADC99091 standard; protein; 609 AA.  
XX  
AC ADC99091;  
XX  
DT 01-JAN-2004 (first entry)  
XX  
DE Human KPP protein - SEQ ID 44.  
XX  
Kw anti-HIV; antiallergic; antiinflammatory; antianaemic; antiparkinsonian;  
Kw neutropic; anticonvulsant; antiarteriosclerotic; hepatotropic; dermatological;  
Kw immunosuppressive; antithyroid; cytostatic; hepatotropic; dermatological;  
Kw osteopathic; nephrotropic; antigout; thyromimetic; neuroprotective;  
Kw urothatic; antihelminthic; antiparasitic; antihelminthic; antiparasitic;  
Kw virucide; protozoacide; fungicide; kinase; phosphatase; KPP;  
Kw cell proliferative disorder; atherosclerosis; cirrhosis; hepatitis;  
Kw cancer; developmental; mental retardation; neurological;  
Kw Alzheimer's disease; Parkinson's; autoimmune; inflammatory; Crohn's;  
Kw diabetes mellitus; viral; bacterial; fungal; parasitic; protozoan;  
Kw helminthic infection; transgenic; gene therapy; human; enzyme.  
XX  
OS Homo sapiens.  
XX  
XX WO2003033680-A2.  
XX  
XX 24-APR-2003.  
XX  
XX 17-OCT-2002; 2002WO-US033723.  
XX  
XX 19-OCT-2001; 2001US-0345474P.  
XX  
XX 02-NOV-2001; 2001US-0343910P.  
XX  
XX 13-NOV-2001; 2001US-0333098P.  
XX  
XX 16-NOV-2001; 2001US-0332424P.  
XX  
XX 30-NOV-2001; 2001US-0334289P.  
XX  
XX (INCY-) INCYTE GENOMICS INC.  
XX  
XX Bandnan O, Baughn WR, Becha SD, Borowsky ML, Duggan BM;  
XX Emerling BM, Forsythe IJ, Gandhi AR, Gorvad AE, Griffin JA;  
XX Gururajan R, Hafalia AFA, Khan FA, Lal PG, Lee EA, Lee SY;  
XX Lindquist EA, Lu DAM, Lu Y, Marquis JP, Nguyen DB, Arvizu CS;  
XX Ramkumar J, Recipon SA, Richardson TW, Swarnakar A, Tang YT;  
XX Thornton MB, Tran UK, Chawla NK, Warren BA, Yang J, Yao MG, Yue H;  
XX Zebarjadian Y;  
XX  
XX WPI; 2003-403214/38.  
XX  
XX N-PSDB; ADC99143.  
XX

PT New human kinases and phosphatases and polynucleotides, useful for  
PT diagnosing, treating or preventing autoimmune or inflammatory disorders  
PT (e.g. AIDS, allergy or anemia), multiple sclerosis, osteoarthritis,  
XX cancer or hepatitis.  
PS Claim 1; SEQ ID NO 44; 424pp; English.  
XX  
CC The invention relates to a novel isolated polypeptide which is a human  
CC kinase and phosphatase (KPP). The KPP polypeptides, polynucleotides,  
CC agonists and antagonists are useful for diagnosing, treating or  
CC preventing cell proliferative disorders such as atherosclerosis,  
CC cirrhosis, hepatitis and cancer, developmental disorders e.g. mental  
CC retardation, neurological disorders including Alzheimer's disease and  
CC Parkinson's disease, autoimmune and inflammatory disorders such as  
CC Crohn's disease and diabetes mellitus and finally, viral, bacterial,  
CC fungal, parasitic, protozoan or helminthic infections. Furthermore, the  
CC polynucleotides encoding KPP may be useful for creating transgenic  
CC animals to model human disease, as well as during gene therapy  
CC procedures. The current sequence is that of the human KPP protein of the  
CC invention.  
XX  
SQ Sequence 609 AA;  
Query Match 87.2%; Score 2182; DB 7; Length 609;  
Best Local Similarity 92.3%; Pred. No. 1.5e-159;  
Matches 430; Conservative 0; Mismatches 0; Indels 36; Gaps 1;  
Qy 1 MLKFOEAAKCVSGSTAISTYPTKTLIARRYVLQQLKSGSGFTYVLSVSDKAKRGEELKVL 60  
Db 1 MLKFOEAAKCVSGSTAISTYPTKTLIARRYVLQQLKSGSGFTYVLSVSDKAKRGEELKVL 60  
Qy 61 KEISVGEINPNETVOANLEAQLSKLDHPAIVKPHASFEQDNFCIITECEGRDLDDKI 120  
Db 61 KEISVGEINPNETVOANLEAQLSKLDHPAIVKPHASFEQDNFCIITECEGRDLDDKI 120  
Qy 121 QYKQAGKIPFNQIIEWFIQLLLGVDMYHERRILHRDLKSKNVFLKNNLLKIGDFGVS 180  
Db 121 QYKQAGKIPFNQIIEWFIQLLLGVDMYHERRILHRDLKSKNVFLKNNLLKIGDFGVS 180  
Qy 181 LLMGSCDLATLTGTPTPHMSPEALKHQGYDTKSDIWSLACILYEMCCMHAFAGSNFLSI 240  
Db 181 LLMGSCDLATLTGTPTPHMSPEALKHQGYDTKSDIWSLACILYEMCCMHAFAGSNFLSI 240  
Qy 241 VLKIVEGDTPSLPERYKELNAMESMLNKNPSLRPSAIEILKIPYLDQQLNLMCRYS 300  
Db 241 VLKIVEGDTPSLPERYKELNAMESMLNKNPSLRPSAIEILKIPYLDQQLNLMCRYS 300  
Qy 301 MTLEDKNLDCQKEAAHIINAMQKRIHLQTLRALSEVQKTPRERMLRKLQAADEKARKL 360  
Db 301 MTLEDKNLDCQKEAAHIINAMQKRIHLQTLRALSEVQKTPRERMLRKLQAADEKARKL 360  
Qy 361 KKIIVEKYENSKEMLSRNFQQLSDVVLHEKTHLKGMEKEQPEGRSLSCSPODEDE 420  
Db 361 KKIIVEKYENSKEMLSRNFQQLSDVVLHEKTHLKGMEKEQPEGRSLSCSPODEDE 420  
Qy 421 ERWQGREESDEPTLENLPESQIPSMDLHELESIVEDATSDLGYN 466  
Db 393 -----ESDEPTLENLPESQIPSMDLHELESIVEDATSDLGYN 430

RESULT 7  
ADC99090  
ID ADC99090 standard; protein; 540 AA.  
XX  
XX ADC99090;  
XX  
XX 01-JAN-2004 (first entry)  
XX  
XX Human KPP protein - SEQ ID 43.  
XX  
XX anti-HIV; antiallergic; antiinflammatory; antianaemic; antiparkinsonian;  
XX neutropic; anticonvulsant; antiarteriosclerotic; antiasthmatic;  
XX immunosuppressive; antithyroid; cytostatic; hepatotropic; dermatological;



CC identifying expressed genes or for physical mapping of human genome. The  
CC encoded polypeptides (ABP68902-ABP69849) are useful as molecular weight  
CC markers, as a food supplement, for generating antibodies, in medical  
CC imaging, screening and diagnostic assays and for treating cell  
CC proliferative disorders (cancer), neurodegenerative diseases (Parkinson's  
CC or Alzheimer's disease), autoimmune diseases (multiple sclerosis,  
CC diabetes, lupus) genetic disorders, myeloid or lymphoid disorders,  
CC platelet or coagulation disorders, wound, burns, incision, ulcers, liver  
CC or lung fibrosis, infections (bacterial, viral, fungal, parasitic),  
CC arthritis, etc. Note: The sequence data for this patent did not form part  
CC of the printed specification, but was obtained in electronic format  
CC directly from WIPO at ftp.wipo.int/pub/published\_pct\_sequences

XX Sequence 497 AA;

Query Match 66.1%; Score 1652; DB 5; Length 497;  
Best Local Similarity 100.0%; Pred. No. 8.8e-119; Mismatches 0; Indels 0; Gaps 0;  
Matches 318; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 149 MHERRILHDLKSNVFLKNNLLKIGDFGVSRLLMGSCDLATLTGTTHYMSPEALKEQG 208  
DB 1 MHERRILHDLKSNVFLKNNLLKIGDFGVSRLLMGSCDLATLTGTTHYMSPEALKEQG 60  
QY 209 YPTKSDIWSLACILYEMCCMNAHAFAGSNFLSIVLXIVGDTSPSLPERYKELNATMESNL 268  
DB 61 YPTKSDIWSLACILYEMCCMNAHAFAGSNFLSIVLXIVGDTSPSLPERYKELNATMESNL 120  
QY 269 NKNPSLRPSAIBILKIPYLDEQLNLMCRYSMTLEDKNLDCQKEAAHI INAMQKRIHLQ 328  
DB 121 NKNPSLRPSAIBILKIPYLDEQLNLMCRYSMTLEDKNLDCQKEAAHI INAMQKRIHLQ 180  
QY 329 TLRALSEVQKMTPRERMRLKLOADEKARKLKKTVEEKYEENSRMQLSRNPFQQLSV 388  
DB 181 TLRALSEVQKMTPRERMRLKLOADEKARKLKKTVEEKYEENSRMQLSRNPFQQLSV 240  
QY 389 DVLHEKTHLKGMEKEQEGRLSCSPQDEDEERWQGREESDEPTLENLPESQIPSPMD 448  
DB 241 DVLHEKTHLKGMEKEQEGRLSCSPQDEDEERWQGREESDEPTLENLPESQIPSPMD 300  
QY 449 LHELESIVEDATSDLYGH 466  
DB 301 LHELESIVEDATSDLYGH 318

RESULT 9

ABB72299  
ID ABB72299 standard; protein; 467 AA.

XX ABB72299;

XX 04-APR-2002 (first entry)

DE Rat protein isolated from skin cells SEQ ID NO: 513.

XX Human; rat; mouse; skin cell; skin wound; cancer; growth defect;  
KW developmental defect; inflammatory disease; dermatological; vulnary;  
KW immunomodulator; anti-inflammatory; cytostatic; neuroprotective.

OS Rattus sp.

XX WO200190357-A1.

PN 29-NOV-2001.

XX 24-MAY-2001; 2001WO-NZ000099.

PF 24-MAY-2000; 2000US-0206650P.

PR 25-JUL-2000; 2000US-0221232P.

XX (GENE-) GENESIS RES & DEV CORP LTD.

PA Watson JD, Strachan L, Sleeman M, Onrust R, Murison JG;  
PI Kumble KD;

XX WPI, 2002-122020/16.

XX New polynucleotides and polypeptides encoded by the polynucleotides  
PT isolated from skin cells, useful for treating skin wounds, cancers,  
PT growth and developmental defects, inflammatory diseases, or for  
PT modulating immune responses.

PS Claim 4; Page 318-319; 466pp; English.

XX The present invention provides the protein and coding sequences of cDNAs  
CC isolated from human, murine and rat skin cell libraries. The sequences  
CC can be used in the development of therapeutic agents useful in the  
CC treatment of skin diseases, including skin wounds, cancer, growth  
CC defects, developmental defects and inflammatory diseases. The proteins  
CC have important roles in the induction of hair growth, cell proliferation  
CC and cell-cell interaction, in maintaining tissue integrity, in wound  
CC healing and in modulating immune responses. The present sequence is a  
CC polypeptide of the invention

XX Sequence 467 AA;

Query Match 59.6%; Score 1491; DB 5; Length 467;  
Best Local Similarity 64.3%; Pred. No. 2.1e-106; Mismatches 40; Indels 80; Gaps 5;  
Matches 301; Conservative 40; Mismatches 47; Indels 80; Gaps 5;  
QY 1 MLKFQEAACVSGSTASTYPTLIARRYVLOOKLGGSGFTVLYVSDKKARGEELKVL 60  
DB 1 MLKFQEAACVSGSTASTYPTLIARRYVLOOKLGGSGFTVLYVSDKKARGEELKVL 58  
QY 61 KEISVGEINNETVQANLEAQLSKLDHPAIVKFAHAFVEQDNFCIITEYCEGRDLDKI 120  
DB 59 KEISVGEINNETVQANLEAQLSKLDHPAIVKFAHAFVEQDNFCIITEYCEGRDLDKI 118  
QY 121 QEYKQAGKIPENCIIEWFQILLGVDMYHERRILHDLKSNVFLKNNLLKIGDFGVSR 180  
DB 119 QEYKQAGKIPENCIIEWFQILLGVDMYHERRILHDLKSNVFLKNNLLKIGDFGVSR 178  
QY 181 LLMGSCDLATLTGTTHYMSPEALKEQGDTKSDIWSLACILYEMCCMNAHAFAGSNFLSI 240  
DB 179 LLMGSCDLATLTGTTHYMSPEALKEQGDTKSDIWSLACILYEMCCMNAHAFAGSNFLSI 238  
QY 241 VLKIVEGDTSPSLPERYKELNATMESMLNKNPSLRPSAIBILKIPYLDEQLNLMCRYS 300  
DB 239 VLKIVEGDTSPSLPERYKELNATMESMLNKNPSLRPSAIBILKIPYLDEQLNLMCRYS 298  
QY 301 MTLEDK-NLDCQKEAAHI INAMQKRIHLQTLRALSEVQKMTPRERMRLKLOADEKARK 359  
DB 299 ATLEDKNSACQKEAAHVNAI----- 320  
QY 360 LKXIVE-EKYEENSRMQLSRNPFQQLSVLXIVGDTSPSLPERYKELNATMESMLNKNPSLRPSAIBILKIPYLDEQLNLMCRYS 418  
DB 321 --KITEGRRYKNNKRAKELRSQNFSGVSAHVL----- 351  
QY 419 DEERWQGREESDEPTLENLPESQIPSPMDLHELESIVEDATSDLYGH 466  
DB 352 -----QELDELTLLESLSQPSLPCNLNDELPSLEGTIVDLGHY 390

RESULT 10

ADCS1182

ID ADCS1182 standard; protein; 357 AA.

XX ADCS1182;

AC ADCS1182;

DT 18-DEC-2003 (first entry)

XX Human cell-cycle related protein, SEQ ID 8.

XX Human; cytostatic; cell-cycle related protein; nuclear export;  
KW nuclear-cytoplasm transport; cytotoxic; cell-cycle control;  
KW immunological disease; neurological disease; cancer.



Db 181 KPNYKSDVWALGCCVYEMATLKHAFAKMDNSLVTRIIEGKLPAMPDRDYSPELAELIRT 240  
QY 267 MLNKNPSLRPSAIEILKIPYLDEQLQNLRCRYSEMPTLED--KNLDCQ-KEAAHIIN---- 319  
Db 241 MLKSRPEERPSVRSILRQPIKQISFFLEATKIKTSKNKKNKNGDSQSPFFATVVSGERAE 300  
QY 320 AMQKRIHLQTLRA-----LSEVQKMTPRERVLRLX--QAADKARKLKKIVEEKEEN 371  
Db 301 SNHEVHPQLPSEGSQTYMGEKCLSOEKPRASGLLKSASLKAHTCKQDL-----SN 355  
QY 372 SKRMOELSRNFQOLSVDVLHETKLGME-----KEEQPE-----GRLSCSPQD 417  
Db 356 TTELATISSVN-----IDIL-----PAKGRDSVSDGFVQENQPRYLDASNELGGI-CSISQ 405  
QY 418 EDEERWQGREESDEPTLENLPESQIPSMDL 449  
Db 406 VEEMLDQNTKSAQP--ENL---IPMWSSDI 432

RESULT 12  
ADB37570  
ID ADB37570 standard; protein; 841 AA.  
AC ADB37570;  
DT 04-DEC-2003 (first entry)  
DE Neural thread protein-related protein #33.  
XX Cytostatic; Antitumour; Antipsoriatic; Dermatological;  
KW Antithrombotic; Antiarteriosclerotic; Vasotropic; Antiinflammatory;  
KW Immunosuppressive; Tranquillizer; Antileptic; Virucide; AD7c-NTP;  
KW Neural thread protein; neuritic sprouting.  
XX Unidentified.  
XX WO2003008444-A2.  
FN 30-JAN-2003.  
XX 19-JUL-2002; 2002WO-CA001106.  
XX 19-JUL-2001; 2001US-0306150P.  
PR 19-JUL-2001; 2001US-0306161P.  
PR 16-NOV-2001; 2001US-0331477P.  
XX (NYWO-) NYMOX CORP.  
XX Averbach PA, Gemmell J;  
PI WPI; 2003-248000/24.  
DR Novel Related peptide or AD7c-neural thread peptide, useful for treating  
PT unwanted cellular proliferations, glandular hyperplasia, unwanted facial  
PT hair, warts and unwanted fatty tissue.  
XX Claim 1; Page 34; 109pp; English.

CC The present invention relates to AD7c-neural thread protein (NTP) and  
CC related proteins and peptides (I; ADB37528-ADB37641). The sequences are  
CC useful for treating a condition in a patient requiring removal or  
CC destruction of cells. The condition can be selected from benign or  
CC malignant tumour of a tissue, hyperplasia, hypertrophy or overgrowth of a  
CC tissue, virally, bacterially or parasitically altered tissue, or  
CC malformation of a tissue, where the tissue is selected from lung, breast,  
CC stomach, pancreas, prostate, bladder, bone, ovary, skin, kidney, sinus,  
CC colon, intestine, stomach, rectum, oesophagus, heart, spleen, salivary  
CC gland, blood, brain and its coverings, spinal cord and its coverings,  
CC muscle, connective tissue, adrenal, parathyroid, thyroid, uterus, testis,  
CC pituitary, reproductive organs, liver, gall bladder, eye, ear, nose,  
CC throat, tonsils, mouth, lymph nodes and lymphoid tissue. The condition is  
CC preferably tonsillar hypertrophy, prostatic hyperplasia, psoriasis,  
CC eczema, dermatosis, cosmetic modification to a tissue (skin, eye, ear,

CC nose, throat, mouth, muscle, connective, hair or breast tissue), vascular  
CC disease (atherosclerosis or arteriosclerosis), haemorrhoids, varicose  
CC veins, inflammatory disease, autoimmune disease, metabolic disease,  
CC hereditary/genetic disease, traumatic disease or physical injury,  
CC nutritional deficiency disease, infectious disease, amyloid disease,  
CC fibrosis disease, storage disease, congenital malformation, enzyme  
CC deficiency disease, poisoning, intoxication, environmental disease,  
CC radiation disease, endocrine disease, degenerative disease and mechanical  
CC diseases. The peptides are useful for treating unwanted cellular  
CC proliferations, glandular (e.g. prostate) hyperplasia, unwanted facial  
CC hair, warts and unwanted fatty tissue, or for preparing antibodies that  
CC recognize and/or bind to Related proteins, Related peptides or NTP  
CC peptides.

XX Sequence 841 AA;  
SQ Query Match 24.5%; Score 613; DB 7; Length 841;  
Best Local Similarity 34.7%; Pred. No. 2.3e-38;  
Matches 157; Conservative 86; Mismatches 153; Indels 56; Gaps 17;  
QY 29 YVLOQKLGSGFVTVYLVSDKKAARGBELKVLKBEISVGLNPNETVOANLEAQLSKLDH 88  
Db 6 YCYLRVVGKSGYGEVTLV--KHRRDGKQVYIKLNLNRNASSRRRAAEQAQLLSQLKH 62  
QY 89 PAIVKEHASFVEQDNFC-IITEYCEGEDDDKICEYKQAGKIPENQIIEFQLLLGVD 147  
Db 63 PNIVTYKESWEGGDLIYVNGCEGDLRYRKE--QKGLLPENOVVFWFOIAVALQ 120  
QY 148 YMERRILHRDLKSKNVFL-KNNLLKIGFVSRLLMGSCDLATTLTGTPHYMSPEALKH 206  
Db 121 YLHEKHILHRDLKTONVFLRTNIIKVGDLGIARVLENHCDMASTLGTFTVYMSPELFSN 180  
QY 207 QGYDTKSDIWSLACILYEMCCMNAFAGSNFLSVLKVIVGDTPLSPERPYPKELNAMES 266  
Db 181 KPNYKSDVWALGCCVYEMATLKHAFAKMDNSLVTRIIEGKLPAMPDRDYSPELAELIRT 240  
QY 267 MLNKNPSLRPSAIEILKIPYLDEQLQNLRCRYSEMPTLED--KNLDCQ-KEAAHIIN---- 319  
Db 241 MLKSRPEERPSVRSILRQPIKQISFFLEATKIKTSKNKKNKNGDSQSPFFATVVSGERAE 300  
QY 320 AMQKRIHLQTLRA-----LSEVQKMTPRERVLRLX--QAADKARKLKKIVEEKEEN 371  
Db 301 SNHEVHPQLPSEGSQTYMGEKCLSOEKPRASGLLKSASLKAHTCKQDL-----SN 355  
QY 372 SKRMOELSRNFQOLSVDVLHETKLGME-----KEEQPE-----GRLSCSPQD 417  
Db 356 TTELATISSVN-----IDIL-----PAKGRDSVSDGFVQENQPRYLDASNELGGI-CSISQ 405  
QY 418 EDEERWQGREESDEPTLENLPESQIPSMDL 449  
Db 406 VEEMLDQNTKSAQP--ENL---IPMWSSDI 432

RESULT 13  
ABUS4632  
ID ABUS4632 standard; protein; 489 AA.  
XX AC ABUS4632;  
XX DT 03-JUN-2003 (first entry)  
XX DE Human NOVX polypeptide #91.  
XX Human; NOVX; metabolic disorder; cardiomyopathy; diabetes; ASD;  
KW hypertension; congenital heart defect; aortic stenosis; valve disease;  
KW atrial septal defect; atrioventricular canal defect; ductus arteriosus;  
KW pulmonary stenosis; subaortic stenosis; ventricular septal defect; VSD;  
KW tuberculous sclerosis; scleroderma; atherosclerosis; infectious disease;  
KW obesity; anorexia; neurodegenerative disorder; Alzheimer's disease;  
KW Parkinson's disease; immune disorder; haematopoietic disorder;  
XX haemophilia; hypercoagulation; Crohn's disease; cancer.  
OS Homo sapiens.

XX	WO200281498-A2.	
XX	17-OCT-2002.	
XX	03-APR-2002; 2002WO-US010780.	
XX	03-APR-2001; 2001US-0281086P.	
XX	03-APR-2001; 2001US-0281136P.	
XX	03-APR-2001; 2001US-0281863P.	
XX	03-APR-2001; 2001US-0281906P.	
XX	06-APR-2001; 2001US-0282020P.	
XX	10-APR-2001; 2001US-0282930P.	
XX	10-APR-2001; 2001US-0282934P.	
XX	12-APR-2001; 2001US-0283512P.	
XX	13-APR-2001; 2001US-0283710P.	
XX	17-APR-2001; 2001US-0284234P.	
XX	19-APR-2001; 2001US-0285325P.	
XX	20-APR-2001; 2001US-0285381P.	
XX	20-APR-2001; 2001US-0285609P.	
XX	23-APR-2001; 2001US-0285748P.	
XX	23-APR-2001; 2001US-0285890P.	
XX	24-APR-2001; 2001US-0286068P.	
XX	25-APR-2001; 2001US-0286292P.	
XX	27-APR-2001; 2001US-0287213P.	
XX	02-MAY-2001; 2001US-0288257P.	
XX	29-MAY-2001; 2001US-0294164P.	
XX	30-MAY-2001; 2001US-0294484P.	
XX	18-JUN-2001; 2001US-0298952P.	
XX	19-JUN-2001; 2001US-0299237P.	
XX	19-JUN-2001; 2001US-0299276P.	
XX	12-SEP-2001; 2001US-0318750P.	
XX	25-SEP-2001; 2001US-0324800P.	
XX	25-SEP-2001; 2001US-0324802P.	
XX	17-OCT-2001; 2001US-0325684P.	
XX	17-OCT-2001; 2001US-0330143P.	
XX	14-NOV-2001; 2001US-0332131P.	
XX	14-NOV-2001; 2001US-0332240P.	
XX	14-NOV-2001; 2001US-0332779P.	
XX	21-NOV-2001; 2001US-0332115P.	
XX	04-DEC-2001; 2001US-0337621P.	
XX	03-JAN-2002; 2002US-0345783P.	
XX	16-JAN-2002; 2002US-0350251P.	
XX	02-APR-2002; 2002US-00114270.	
XX	(CURA-) CURAGEN CORP.	
XX	Guo X, Kekuda R, Miller CE, Malyankar UM, Szytek KA; Pattarajan M, Liu X, Gusev VY, Li L, Vernet CAM, Zerhusen BD; Gorman L, Shenoy SG, Pena CE, Smithson G, Burgess CE, Gerlach V; Padigaru M, Shinkets RA, Gangolli EA, Taupier RJ, Casman SJ, Ji W; Anderson DW, Leite WJ, Rastelli L, Edinger SR, Stone DJ; Macdougall JR, Rothenberg ME, Mazur A, Millet I, Peyman JA; Ellerman K;	
XX	WPI; 2003-046858/04.	
XX	N-PSDB; ABX72260.	
XX	New isolated NOVX polypeptide useful for treating atherosclerosis, metabolic disorders, diabetes, obesity, infectious disease, anorexia, neurodegenerative disorders, Alzheimer's disease and cancer.	
XX	Claim 1; Page 288; 666pp; English.	
XX	The invention relates to human polypeptides, termed NOVX, and the polynucleotides encoding them. The polypeptides and polynucleotides are useful for diagnosing disease, and screening for potential therapeutic agents. The sequences are useful for treating metabolic disorders, aortic cardiomyopathy, diabetes, hypertension, congenital heart defects, aortic stenosis, atrial septal defect (ASD), atrioventricular canal defect, ductus arteriosus, pulmonary stenosis, subaortic stenosis, ventricular septal defect (VSD), valve diseases, tuberosus sclerosis, scleroderma, atherosclerosis, obesity, infectious disease, anorexia, neurodegenerative	

QY	332	ALSEVQKMTPRERMLRLKLQAADKARKLKKIIVEEKYEENSKRMQBLRSNFOOLSDVL	391
Db	345	---PVKKNSGEERKOWSEEAA--KKRRLFIEKEK-----KQKDQIRFLKAQOMK---R	391
QY	392	HEKTHLKGWEKEKEP-----EGRLS---CSPO-----	416
Db	392	QEQRLEKNRARSGWRNVLRAGSGEVKASFGIGAVSPSCPSRGQYEHYHAIFDQ	451
QY	417	-----DEDERWQ3 425	
Db	452	MQLRAEDNEARWKG 466	
 RESULT 15			
ABP71711			
ID	ABP71711	standard; protein; 774 AA.	
XX	ABP71711;		
XX	AC AC		
DT	17-APR-2003	(first entry)	
DE	NEK-like serine/threonine protein kinase NEK1_mouse # SEQ ID 3.		
KX	NEK-like serine/threonine protein kinase; cytostatic; cardiant; antiinflammatory; nootropic; neuroprotective; cancer; colon cancer; cardiovascular disorder; diabetes; COPD; CNS disorder; mouse; rat.		
OS	Mus musculus.		
OS	Rattus norvegicus.		
PN	WC02030300903-A2.		
PD	03-JAN-2003.		
PF	24-JUN-2002; 2002WO-EP006948.		
XX	25-JUN-2001; 2001US-0300068P.		
XX	07-DEC-2001; 2001US-0336704P.		
PA	(FARB ) BAYER AG.		
PI	Xiao Y;		
XX	WPI: 2003-184051/18		

XX New polynucleotide encoding a NEK-like serine/threonine kinase  
 PT polypeptide useful for treating diseases associated with kinase  
 PT dysfunction, e.g. cardiovascular disorders, cancer such as colon cancer,  
 PT diabetes and CNS disorders.  
 XX  
 PS Disclosure; Fig 3; 149pp; English.  
 XX  
 CC The invention relates to a newly isolated polynucleotide encoding an NEK-  
 CC like serine/threonine protein kinase. The activity of the polynucleotide  
 CC and polypeptide of the invention may be described as cytostatic,  
 CC cardiant, antiinflammatory, nootropic and neuroprotective. The expression  
 CC vector and reagent of the invention are useful for the preparation of a  
 CC medicament for modulating the activity of an NEK-like serine/threonine  
 CC kinase in a disease, such as cancer (e.g. colon cancer), cardiovascular  
 CC disorder, diabetes, COPD or CNS disorder. The polypeptides may also be  
 CC used to identify compounds which may act as activators or inhibitors at  
 CC the enzyme's active site, to raise specific antibodies which can block  
 CC the enzyme and effectively reduce its activity, as a bait protein in a  
 CC two-hybrid or three-hybrid assay to identify other proteins which bind to  
 CC or interact with the human NEK-like serine/threonine kinase polypeptide  
 CC and modulate its activity, and for the immunisation of mammals. The  
 CC current sequence represents an NEK-like serine/threonine protein kinase  
 CC designated NEK1\_mouse  
 XX  
 CC Sequence 774 AA;  
 XX  
 SQ Query Match 23.9%; Score 598; DB 6; Length 774;

Query Match 23.9%; Score 598; DB 6; Length 774;  
Best Local Similarity 28.5%; Pred. No. 3e-37;



us-10-803-278-4.rag

Mon Oct 4 10:35:28 2004

```

Matches 141; Conservative 94; Mismatches 132; Indels 128; Gaps 13;
QY 28 RYVLOKLGSGFQTYVYSDKAKRGEBELKVLKEISVGEINPNETVQANLEAQLLSKLD 87
Db :||:||||| :||:||||| :||:||||| :||:||||| :||:||||| :||:|||||
3 KTVRLQIGEGSGFQAVLV--KSTEDGRHY-VIKENISRMSDKERQSRREVAVLANMK 59
QY 88 HPAIVKHFASVFEODNFCIITEYCEGRDLDKIOBYKQAGKIFPENQIIEWFIQLLQVD 147
Db :||:||||| :||:||||| :||:||||| :||:||||| :||:||||| :||:|||||
60 HFNIVQYKESFENGSLYIWMYCEGDLFKIN--AQKGALFQEDQILDNMFVQICLALK 117
QY 148 YMHERRILHRDLKSKNVFL-KNNLLKIGDFGVSRLLMGSCDLATTLTGTPHYMSPEALKH 206
Db :||:||||| :||:||||| :||:||||| :||:||||| :||:||||| :||:|||||
118 HVHDKILHRDIKSNIFLTQDGTGVLGDFGIARVLNSTVELARTICIGTPYLSPEICEN 177
QY 207 QGYDTKSDIWSLACILYEMCCNHFAGSNFLSIVLKIVEGDTPSLPERYPKELNAINES 266
Db :||:||||| :||:||||| :||:||||| :||:||||| :||:||||| :||:|||||
178 KPYNNKSDIWAUGCVLYELCTLKHAFAGNMKNLVLLKIIISGSFPVPSPHYSDYDLRSLLSQ 237
QY 267 MLNKNPSLRPSAIEILKIPYLDLQNLN-----C----- 296
Db :||:||||| :||:||||| :||:||||| :||:||||| :||:||||| :||:|||||
238 LFKRNRDRPSVNSILEKGFIAKRIEKFLSPOLIAEFCFKLTKSKFGPOPLPGKRPASQ 297
QY 297 -----RYSEWTLKNDLDCQKEAAHIIINAMQKRIHLQTLR 331
Db :||:||||| :||:||||| :||:||||| :||:||||| :||:||||| :||:|||||
298 GVSSFVPAQKITKPAAYGVPLTYKYGDKKLEKPPPKHQAHOI----- 344
QY 332 ALSEVQKMTPRMRRLKLADEKARKLKIVEEKYENSKRMOELSRNFQQLSVDVL 391
Db :||:||||| :||:||||| :||:||||| :||:||||| :||:||||| :||:|||||
345 ---PVKMNSGEERKMSSEEA--KKRLEFIEKEK-----KQDQIRFLKAEQMK--R 391
QY 392 HEXTHLKGMEEKEEQP-----EGRLS---CSPQ----- 416
Db :||:||||| :||:||||| :||:||||| :||:||||| :||:||||| :||:|||||
392 QEQRLERINRAREOGWRNVLRAGSGSEVKASFFGTGGAVSPSPSPGQYEHYHAFDQ 451
QY 417 -----DEDERWQG 425
Db :||:||||| :||:||||| :||:||||| :||:||||| :||:||||| :||:|||||
452 MQLRAEDNEARWKG 466

```

Search completed: October 1, 2004, 11:18:07  
Job time : 130 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 1, 2004, 11:16:00 ; Search time 33 Seconds  
(without alignments)  
754,052 Million cell updates/sec

Title: US-10-803-278-4  
Perfect score: 2501  
Sequence: 1 MLKFEAAKCVSGSTAISTY.....LQYHGDCLNLSLDYWKNEK 482

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 389414 seqs, 51625971 residues

Total number of hits satisfying chosen parameters: 389414

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database :  
1: /cgn2\_6/ptodata/2/iaa/5A\_COMB.pdp:\*  
2: /cgn2\_6/ptodata/2/iaa/5B\_COMB.pdp:\*  
3: /cgn2\_6/ptodata/2/iaa/6A\_COMB.pdp:\*  
4: /cgn2\_6/ptodata/2/iaa/6B\_COMB.pdp:\*  
5: /cgn2\_6/ptodata/2/iaa/PTUS\_COMB.pdp:\*  
6: /cgn2\_6/ptodata/2/iaa/backfiles1.pdp:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	598	23.9	616	4	US-09-873-404-4
2	584	23.4	640	4	US-09-873-404-2
3	495	19.8	445	4	US-09-167-206-2
4	493	19.7	345	3	US-09-173-581-6
5	493	19.7	345	3	US-09-420-915-6
6	454	18.2	692	4	US-09-992-481-2
7	445	17.8	302	4	US-09-579-6648-9
8	443	17.7	302	3	US-09-221-235-2
9	443	17.7	302	3	US-09-221-928-2
10	443	17.7	302	3	US-09-221-527-2
11	443	17.7	302	3	US-09-221-236-2
12	443	17.7	302	3	US-09-221-416-2
13	443	17.7	302	3	US-09-221-245-2
14	443	17.7	302	3	US-09-163-115-2
15	443	17.7	302	3	US-09-221-528-2
16	443	17.7	302	3	US-09-593-553-2
17	443	17.7	302	3	US-09-221-237-2
18	440.5	17.6	699	4	US-09-457-0408-18
19	429.5	17.2	979	3	US-08-870-529-2
20	429.5	17.2	979	4	US-09-544-794-2
21	404.5	16.2	233	2	US-08-712-709-1
22	404.5	16.2	233	3	US-09-111-444-1
23	404.5	16.2	233	3	US-09-541-228-1
24	395.5	15.8	293	3	US-09-049-671-1
25	395.5	15.8	293	3	US-09-295-068-1
26	382.5	15.3	239	2	US-08-712-709-7
27	382.5	15.3	239	2	US-09-049-671-3

US-09-873-404-4  
; Sequence 4, Application US/09873404  
; Patent No. 650656  
; GENERAL INFORMATION:  
; APPLICANT: WEBSTER, Marion et al  
; TITLE OF INVENTION: ISOLATED HUMAN KINASE PROTEINS, NUCLEIC  
; TITLE OF INVENTION: ACID MOLECULES ENCODING HUMAN KINASE PROTEINS, AND USES  
; TITLE OF INVENTION: THEREOF  
; FILE REFERENCE: CL001212-CIP  
; CURRENT APPLICATION NUMBER: US/09/873,404  
; CURRENT FILING DATE: 2001-06-05  
; NUMBER OF SEQ ID NOS: 4  
; SOFTWARE: FastSEQ for Windows Version 4.0  
; SEQ ID NO 4  
; LENGTH: 616  
; TYPE: PRT  
; ORGANISM: Mouse  
US-09-873-404-4

RESULT 1  
US-09-873-404-4  
; Sequence 4, Application US/09873404  
; Patent No. 650656  
; GENERAL INFORMATION:  
; APPLICANT: WEBSTER, Marion et al  
; TITLE OF INVENTION: ISOLATED HUMAN KINASE PROTEINS, NUCLEIC  
; TITLE OF INVENTION: ACID MOLECULES ENCODING HUMAN KINASE PROTEINS, AND USES  
; TITLE OF INVENTION: THEREOF  
; FILE REFERENCE: CL001212-CIP  
; CURRENT APPLICATION NUMBER: US/09/873,404  
; CURRENT FILING DATE: 2001-06-05  
; NUMBER OF SEQ ID NOS: 4  
; SOFTWARE: FastSEQ for Windows Version 4.0  
; SEQ ID NO 4  
; LENGTH: 616  
; TYPE: PRT  
; ORGANISM: Mouse  
US-09-873-404-4

ALIGNMENTS

US-09-873-404-4  
; Sequence 4, Application US/09873404  
; Patent No. 650656  
; GENERAL INFORMATION:  
; APPLICANT: WEBSTER, Marion et al  
; TITLE OF INVENTION: ISOLATED HUMAN KINASE PROTEINS, NUCLEIC  
; TITLE OF INVENTION: ACID MOLECULES ENCODING HUMAN KINASE PROTEINS, AND USES  
; TITLE OF INVENTION: THEREOF  
; FILE REFERENCE: CL001212-CIP  
; CURRENT APPLICATION NUMBER: US/09/873,404  
; CURRENT FILING DATE: 2001-06-05  
; NUMBER OF SEQ ID NOS: 4  
; SOFTWARE: FastSEQ for Windows Version 4.0  
; SEQ ID NO 4  
; LENGTH: 616  
; TYPE: PRT  
; ORGANISM: Mouse  
US-09-873-404-4

Qy	28	RYVLQKLGSGSGTGYLVSDKAKGEELKVLKISVGEINPNETVQANLEAQLSKLD	87
Db	3	KYVRLQKIGSGSGKAVLV--KSTEDGRHY-VIKEINISRMXSKRQESRREVAVLANK	59
Qy	88	HPATVKEFHAFVQDNFCITTEYCEGRDLDDKIQEVKQAGKIFPENQIIEFWIQLLIGVD	147
Db	60	HPNIVQKESFEENGSLYIVNDYCEGGDLFKRIN--AQKGAIFQEDQILDFVQICLAK	117
Qy	148	YMHERRILHRLDLKSNVFL-KNNLLKIGDFGVSRLLMGSCDLATTITGTTHYMSPEALKH	206
Db	118	HVHEDRKILHRDIKSNIFLTQDGTQVQLGDFGIARVLNSTVELARTCTGTPTYSPEICEN	177
Qy	207	QGVDTKSDIWSLACILYEMCCMHAFAGSNFLSVLKVIGEDTSPSLPERYPKELNAMES	266
Db	178	KPNYKSDIWSLACILYEMCCMHAFAGSNFLSVLKVIGEDTSPSLPERYPKELNAMES	237
Qy	267	MLKNPRLRPSAIEILKIPYLDQLQVLM-----C-----	296
Db	238	LFXENPRDRPSVNSILEKGFIAKRIFLSPQIAAEFCLTKLSKFGPQLPKGRPASGQ	297
Qy	297	-----RYSENTLEDKNDCKEAAHINAMQKRIHLQTLR	331
Db	298	GVSSFVPAQKITPAKYGVPLTYKYGDKKLEKPPPKHQAHOI-----	344
Qy	332	ALSEVQMTPRMRRLKQADEKARKKKIVVEEYKENSKEMLRSNFQOLSVDVL	391

```
Db 345 ---PVKKNKSGBEKKNVSEAA--KKRLLEFTEKE-----KQDQIRFLKAEQMK---R 391
QY 392 HEKTHLKGEKEEQP-----EGRLS---CSPO----- 416
Db 392 QEQRLERINRAEQWRNVLRAAGSGEVKASFFGIGGAVSPSPCSGPGQYEHYHAIFDQ 451
QY 417 -----DEDEERWQG 425
Db 452 MORLRAEDNEARWKG 466

RESULT 2
US-09-873-404-2
; Sequence 2, Application US/09873404
; Patent No. 6500656
; GENERAL INFORMATION:
; APPLICANT: WEBSTER, Marion et al
; TITLE OF INVENTION: ISOLATED HUMAN KINASE PROTEINS, NUCLEIC
; TITLE OF INVENTION: ACID MOLECULES ENCODING HUMAN KINASE PROTEINS, AND USES
; TITLE OF INVENTION: THEREOF
; FILE REFERENCE: CL001212-CIP
; CURRENT APPLICATION NUMBER: US/09/873,404
; CURRENT FILING DATE: 2001-06-05
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 640
; TYPE: PRT
; ORGANISM: Human
US-09-873-404-2

Query Match 23.4%; Score 584; DB 4; Length 640;
Best Local Similarity 29.2%; Pred. No. 5.8e-38;
Matches 149; Conservative 94; Mismatches 164; Indels 104; Gaps 14;

QY 28 RYVLOQKLGSGFGTVLVSDKKAKGEBELKVLKEISVGEINPNETVQANLEAQLLSKLD 87
Db 3 KYDVIKAIQGAAGFKAYLA---KGKSDSKHCVCVKEINPEKMPIQEAKSKKEVILLERKM 59
QY 88 HPAIVKFAHAFVEQDNFCITTEYCEGRLDDKIQEYKQAGKIFPENQIIEWFIQLLQVD 147
Db 60 HPNIVAFNFSQNGRLFTVMEYCDGGLMKRN--RQGVLFSDQILGWFPVQISGLK 117
QY 148 YHERILHRDLKSKNVFLKN--LLKIGDFGVSRLLMGSCDLATTLTGTPHYMSPEALK 205
Db 118 HIHDKILHRDIKAQNIFLSKNGWAKLGFGLIARVLNNSMELARTCIGTPYLSPEICQ 177
QY 206 HQGYDTPKSDIWSLACILYEMCMNHAFAGSNFLSVILKIVEGDTPSLPERYPKELNAIWE 265
Db 178 NKPNYNTDWSLGVLYELCTLKHPFEGNNLQQLVLKICQAHFAPISPGFSRELHSLIS 237
QY 266 SWLNKPNLRPSAIEILKIPYLDLQQLNLMCRYSEMTELDKNDLDCQKAAHII----- 318
Db 238 QLFQVSPDRDPSINSILKRPF---LENLIPKY--LTPE-----VQEEFSEHMLICEAGAP 287
QY 319 ---NAMQKRIHLQTRALSEVQKQTPRERMLR-----KLOAADEKARKLK 361
Db 288 ASRHAGKVVQKQIKVFRFG---KCPPRSISVPIKRNAILHRNEWRPPAGAKARSIGK 344
QY 362 KIVEEK-----YEENSKRMQELSRNFQQLSDVVLHEKTHLKG----- 399
Db 345 MIERPKIAAVCGHYDYVYLAQLMLRRRAHKPSVHIPQENTGVEDYQGTNRGPGSPQWP 404
QY 400 ---MEEKEEQEGRLSCSPQ-----DEDEERWQREE---ESDEPTLENLPESQPIPSMD 448
Db 405 AEYLQKFEAQQVKLVKQGLRPPSSAEPNPNQRELNSGEBRPFQELPRK----- 458
QY 449 LHELESIVEDATSDLGWGDGNCNLSIDENYK 479
Db 459 -----NEMKQOEYWK 468

RESULT 3
```

```
US-09-167-206-2
; Sequence 2, Application US/09167206A
; Patent No. 6476193
; GENERAL INFORMATION:
; APPLICANT: Nandabalan, Krishnan
; APPLICANT: Schulz, Vincent P.
; APPLICANT: Yang, Mei-Jia
; TITLE OF INVENTION: NIK1 PROTEIN AND NIK1 PROTEIN COMPLEXES
; FILE REFERENCE: 15966-521 NIK1 protein complexes
; CURRENT APPLICATION NUMBER: US/09/167,206A
; CURRENT FILING DATE: 1998-10-06
; NUMBER OF SEQ ID NOS: 26
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 2
; LENGTH: 445
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-167-206-2

Query Match 19.8%; Score 495; DB 4; Length 445;
Best Local Similarity 32.7%; Pred. No. 3.8e-31;
Matches 139; Conservative 70; Mismatches 168; Indels 48; Gaps 14;

QY 26 ARRYVLQKLGSGFGTVLVSDKKAKGEBELKVL--KEISVGEINPNETVQANLEAQLL 83
Db 5 AEDYEVLYTIGTSGYGRG-----QKIRKSDGKILVWKELDYGSMTAEKQMLVSEVNL 59
QY 84 SKLDHPAIVKFAHAFVEQDN--FCITTEYCEGRLDDKIQEYKQAGKIFPENQIIEWFIQ 141
Db 60 RELKHPNIVRYDRIIDRTNTTLYIVMEYCEGDLASVITKTKERQYLDLDEFVLKWTQ 119
QY 142 LLGLGVDMYHER-----RIHRDLKSKNVFL--KQNLKIGDFGVSRLLMGSCDLATTLTG 194
Db 120 LTLALKECHERSDGGHTVLHRDLKPNVFLDGQN--VKLGDFGLARILNHDTSFAKTFVG 178
QY 195 TPHYMSPEALKHQGYDTPKSDIWSLACILYEMCMNHAFAGSNFLSVILKIVEGDTPSLPE 254
Db 179 TPYMSPEQNNRNSYNEKSDIWSLGLLYELCALMPPTAFSQKELAGKIREGKFRIRPY 238
QY 255 RYPKELNAIMESMLNKNPSLRPSAIEILKIPYL-----DEQLNLMCRYSEMTELDKNDL 309
Db 239 RYSDLENEITRMLNKDHYRPSVEILENPLIADLVADEQRNLERRGQLGEPEKSD 298
QY 310 CQKEAAHIINAMQ--KRIHLQ-----TLRALSE-----VQKQTPRERMLRKLQAADEKARK 359
Db 299 ---SSPVLSELKXELQOERERALKAREERLEQKEQELCVREERLAEDKLARAEENLKN 354
QY 360 LKKIVEKYEENSKRMQELSRNFQQLSDVVLHEKTHLKGMBEKE-----EQEPEGRISCS 414
Db 355 YSLDKERKFLSLASNPPELL-----NLPSSVKKKVFHFG--ESKENIMRSENSESQITSK 407
QY 415 PQDED 419
Db 408 SKCKD 412

RESULT 4
US-09-173-581-6
; Sequence 6, Application US/09173581A
; Patent No. 6013455
; GENERAL INFORMATION:
; APPLICANT: Bandman, Olga
; APPLICANT: Tang, Y. Tcm
; APPLICANT: Hillman, Jennifer L.
; APPLICANT: Yue, Henry
; APPLICANT: Guegler, Karl J.
; APPLICANT: Corley, Neil C.
; APPLICANT: Gorgone, Gina
; APPLICANT: Azimzai, Yalda
; APPLICANT: Lu, Aina
; TITLE OF INVENTION: Protein Kinase Homologs
; FILE REFERENCE: PF-0614 US
; CURRENT APPLICATION NUMBER: US/09/173,581A
```

```

; CURRENT FILING DATE: 1998-10-15
; NUMBER OF SEQ ID NOS: 18
; SOFTWARE: PERL Program
; SEQ ID NO 6
; LENGTH: 345
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE: -
; OTHER INFORMATION: 1567782
US-09-173-581-6

Query Match
Best Local Similarity 19.7%; Score 493; DB 3; Length 345;
Matches 116; Conservative 89; Mismatches 119; Indels 96; Gaps 11;

QY 28 RYVLOQKLGSGFGTGYLVSDKKAKEGELKVLKEISVGEINPNETVQANLEAQLLSKLD 87
Db 3 KYVRLQKIGESFGKAILV--KSTEDGRQY-VIKEINISRMSSKERESRREAVLANMK 59
QY 88 HPAIVKPHASVFEQDNFCIITEYCEGRDLDKIQEYKQAGKIPPENQIIEWFIQLLGV 147
Db 60 HPNIVQYRESF-----EG-----ILDMFVQICLALK 85
QY 148 YMHERRIHLRDLKSNVFL-KNNLLKIGDFGVSRLMGSCDLATLTGTGPHYMSPEALKH 206
Db 86 HVHDKILHRDLKSNVFLKDGTVQLGDFGIARVLNSTVELARTICIGTPYLSPEICEN 145
QY 207 QGYDTKSDIWSLACILYEMCCMHAFAGSNFLSVILKIVEGTPSLPERRYKELNAMES 266
Db 146 KPYNNKSDIWSLACILYEMCCMHAFAGSNFLSVILKIVEGTPSLPERRYKELNAMES 266
QY 267 MLNKNPSLRPSAIEILKIPYLDEQLNLM-----CRYSEMTLEKNDLQCKEAAHI 317
Db 206 LFXENPRDRPSVNSILEKGFIAKRIEFLSPOLIAEFCCLKTSKFSQPIPAKRPASG- 264
QY 318 INAMOKRIHLQTLRALSEVQKMT-PRERMLRKLQAADKARKLKIIVEKYEENSXRMQ 376
Db 265 -----QNSISVMPAQKITKPAKYGI-----PLAYKKYGDKK----- 296
QY 377 ELRSNFQQLSDVVLHKKHKLKGVKEEKEOPEGRILSCSPQDEDERWQGREESDEPTLE 436
Db 297 -----LHEKKPLQKHQAHTPEKRVNTG-----EERRKISEEAAKRRLE 337

RESULT 5
US-09-420-915-6
; Sequence 6, Application US/09420915
; Patent No. 6264947
; GENERAL INFORMATION:
; APPLICANT: Bandman, Olga
; APPLICANT: Tang, Y. Tom
; APPLICANT: Hillman, Jennifer L.
; APPLICANT: Yue, Henry
; APPLICANT: Guegler, Karl J.
; APPLICANT: Corley, Neil C.
; APPLICANT: Gorgone, Gina
; APPLICANT: Azimzai, Valda
; APPLICANT: Lu, Aina
; TITLE OF INVENTION: Protein Kinase Homologs
; FILE REFERENCE: PP-0614 US
; CURRENT APPLICATION NUMBER: US/09/420,915
; EARLIER FILING DATE: 1999-10-20
; EARLIER FILING DATE: 1999-10-20
; EARLIER FILING DATE: 1998-10-15
; NUMBER OF SEQ ID NOS: 18
; SOFTWARE: PERL Program
; SEQ ID NO 6
; LENGTH: 345
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE: -
; OTHER INFORMATION: 1567782
US-09-420-915-6
```

```

Query Match
Best Local Similarity 19.7%; Score 493; DB 3; Length 345;
Matches 116; Conservative 89; Mismatches 119; Indels 96; Gaps 11;

QY 28 RYVLOQKLGSGFGTGYLVSDKKAKEGELKVLKEISVGEINPNETVQANLEAQLLSKLD 87
Db 3 KYVRLQKIGESFGKAILV--KSTEDGRQY-VIKEINISRMSSKERESRREAVLANMK 59
QY 88 HPAIVKPHASVFEQDNFCIITEYCEGRDLDKIQEYKQAGKIPPENQIIEWFIQLLGV 147
Db 60 HPNIVQYRESF-----EG-----ILDMFVQICLALK 85
QY 148 YMHERRIHLRDLKSNVFL-KNNLLKIGDFGVSRLMGSCDLATLTGTGPHYMSPEALKH 206
Db 86 HVHDKILHRDLKSNVFLKDGTVQLGDFGIARVLNSTVELARTICIGTPYLSPEICEN 145
QY 207 QGYDTKSDIWSLACILYEMCCMHAFAGSNFLSVILKIVEGTPSLPERRYKELNAMES 266
Db 146 KPYNNKSDIWSLACILYEMCCMHAFAGSNFLSVILKIVEGTPSLPERRYKELNAMES 266
QY 267 MLNKNPSLRPSAIEILKIPYLDEQLNLM-----CRYSEMTLEKNDLQCKEAAHI 317
Db 206 LFXENPRDRPSVNSILEKGFIAKRIEFLSPOLIAEFCCLKTSKFSQPIPAKRPASG- 264
QY 318 INAMOKRIHLQTLRALSEVQKMT-PRERMLRKLQAADKARKLKIIVEKYEENSXRMQ 376
Db 265 -----QNSISVMPAQKITKPAKYGI-----PLAYKKYGDKK----- 296
QY 377 ELRSNFQQLSDVVLHKKHKLKGVKEEKEOPEGRILSCSPQDEDERWQGREESDEPTLE 436
Db 297 -----LHEKKPLQKHQAHTPEKRVNTG-----EERRKISEEAAKRRLE 337

RESULT 6
US-09-992-481-2
; Sequence 2, Application US/09992481
; Patent No. 6593125
; GENERAL INFORMATION:
; APPLICANT: Fiddle, Carl Johan
; APPLICANT: Hilbun, Erin
; APPLICANT: Mathur, Brian
; APPLICANT: Turner, C. Alexander Jr.
; TITLE OF INVENTION: No. 6593125el Human Kinases and Polynucleotides
; FILE REFERENCE: Encoding the Same
; CURRENT APPLICATION NUMBER: US/09/992,481
; PRIOR FILING DATE: 2001-11-19
; PRIOR FILING DATE: 2000-11-20
; NUMBER OF SEQ ID NOS: 5
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 692
; TYPE: PRT
; ORGANISM: homo sapiens
US-09-992-481-2

Query Match
Best Local Similarity 18.2%; Score 454; DB 4; Length 692;
Matches 97; Conservative 61; Mismatches 103; Indels 8; Gaps 4;

QY 28 RYVLOQKLGSGFGTGYLVSDKKAKEGELKVLKEISVGEINPNETVQANLEAQLLSKLD 87
Db 3 KYERIRVVGGAIGIVHLC---LRKADQKLVIIKQIPVEQMTKEERQAQNECQVLLN 59
QY 88 HPAIVKPHASVFEQDNFCIITEYCEGRDLDKIQEYKQAGKIPPENQIIEWFIQLLGV 147
Db 60 HPNVIEYENFLEDKALMIAMEYAPGGTIAEFTQ--KRCNSLLEESTILHFFVQILLALH 117
QY 148 YMHERRIHLRDLKSNVFLKNN--LLKIGDFGVSRLMGSCDLATLTGTGPHYMSPEALK 205
Db 118 HVHDLILHRDLKSNVFLKNN--LLKIGDFGVSRLMGSCDLATLTGTGPHYMSPEALK 176
```

QY 206 HQGYDTSKDIWSLACILYEMCCMHAFAGSNFLSIVLKIIVEGDTSPSLPERYPKELNALME 265  
DB 177 GRPYNGKSDIWAIGCVLYELASLAKRAFAANLPAFLVKIMSGTFAPISTRYSPELRQLVL 236  
QY 266 SMLNKNPSLRPSAIELTKIPYDEQLNL 294  
DB 237 SLSLSEPAQRPPPLSHIMAQPLCIRALLNL 265

RESULT 7  
US-09-579-664B-9  
; Sequence 9, Application US/09579664B  
; Patent No. 654719  
; GENERAL INFORMATION:  
; APPLICANT: Immunex Corporation  
; APPLICANT: Bird, Timothy A.  
; APPLICANT: Virca, G. Duke  
; APPLICANT: Martin, Unja  
; APPLICANT: Anderson, Dirk M.  
; TITLE OF INVENTION: NOVEL MURINE AND HUMAN KINASES  
; FILE REFERENCE: 2923-A  
; CURRENT APPLICATION NUMBER: US/09/579,664B  
; CURRENT FILING DATE: 2000-05-26  
; NUMBER OF SEQ ID NOS: 36  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 9  
; LENGTH: 302  
; TYPE: PRT  
; ORGANISM: Mus musculus  
US-09-579-664B-9

Query Match 17.8%; Score 445; DB 4; Length 302;  
Best Local Similarity 36.2%; Pred. No. 2e-27;  
Matches 92; Conservative 53; Mismatches 95; Indels 14; Gaps 7;  
QY 31 LOOKLGGSGFGTYLVVS---DKKAKGEBELKVLKEISVGEEL-NPNETVQANLEAQLLSKLDH 86  
DB 36 IEKKIGRGQFSEVYRASCLLD---GVFV-ALKKVOIFDLMDAKARADCIKEIDLKQLNH 89  
QY 87 DHPAIVKPHASVFEQDNFCIITEYCEGRDLDDKIOEYKQAGKIFPENQIIEFWIQLLGV 146  
DB 90 NHPNVIKYASFIEDNELNIVLELADAGDLGRFFSSKTTAAHSLVGTPTVYMSPERIH 149  
QY 147 DYMERRILHRDLKSKNVFL-KNNLLKIGDFGVSRLLMGSCDLATLTGTPTHYMSPEALK 205  
DB 150 DHMSRRVHRDIKPNANVITATGVVKGDLGLGRFFSSKTTAAHSLVGTPTVYMSPERIH 209  
QY 206 HQGYDTSKDIWSLACILYEMCCMHAFAGS--NFLSIVLKIIVEGDTSPSL-ERYPKELNA 262  
DB 210 ENGYNFKSDIWSLGLLYEMAAALQSPFYGDKMNLKIEQCDYPPPLPSDHYSEELRQ 269  
QY 263 IMESMLNKNPSLRP 276  
DB 270 LVNINPDEKRP 283

RESULT 8  
US-09-221-235-2  
; Sequence 2, Application US/09221235  
; Patent No. 6043040  
; GENERAL INFORMATION:  
; APPLICANT: Acton, Susan  
; TITLE OF INVENTION: NOVEL CSAPK-1 NUCLEIC ACID MOLECULES AND USES THEREFOR  
; FILE REFERENCE: WNI-050  
; CURRENT APPLICATION NUMBER: US/09/221,235  
; CURRENT FILING DATE: 1998-12-28  
; EARLIER APPLICATION NUMBER: 09/163,115  
; EARLIER FILING DATE:  
; NUMBER OF SEQ ID NOS: 15  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 2  
; LENGTH: 302

; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-09-221-235-2  
Query Match 17.7%; Score 443; DB 3; Length 302;  
Best Local Similarity 34.6%; Pred. No. 2.9e-27;  
Matches 90; Conservative 58; Mismatches 102; Indels 10; Gaps 6;  
QY 31 LOOKLGGSGFGTYLVSDKKAKGEBELKV-LKEISVGEEL-NPNETVQANLEAQLLSKLDH 88  
DB 36 IEKKIGRGQFSEVY---RAACLLDGVFVALKKVOIFDLMDAKARADCIKEIDLKQLNH 91  
QY 89 PAIVKPHASVFEQDNFCIITEYCEGRDLDDKIOEYKQAGKIFPENQIIEFWIQLLGV 148  
DB 92 PNVIKYASFIEDNELNIVLELADAGDLGRFFSSKTTAAHSLVGTPTVYMSPERIH 151  
QY 149 MHERRILHRDLKSKNVFL-KNNLLKIGDFGVSRLLMGSCDLATLTGTPTHYMSPEALKHQ 207  
DB 152 MHSRRVHRDIKPNANVITATGVVKGDLGLGRFFSSKTTAAHSLVGTPTVYMSPERIH 211  
QY 208 GYDTSKDIWSLACILYEMCCMHAFAGS--NFLSIVLKIIVEGDTSPSL-ERYPKELNALME 264  
DB 212 GYNFKSDIWSLGLLYEMAAALQSPFYGDKMNLKIEQCDYPPPLPSDHYSEELRQ 271  
QY 265 ESMLNKNPSLRPSAIELKI 284  
DB 272 NMCINPDEKRPDVTYVDV 291

RESULT 9  
US-09-221-928-2  
; Sequence 2, Application US/09221928  
; Patent No. 6121030  
; GENERAL INFORMATION:  
; APPLICANT: Acton, Susan  
; TITLE OF INVENTION: NOVEL CSAPK-1 NUCLEIC ACID MOLECULES AND USES THEREFOR  
; FILE REFERENCE: WNI-050  
; CURRENT APPLICATION NUMBER: US/09/221,928  
; CURRENT FILING DATE: 1998-12-28  
; EARLIER APPLICATION NUMBER: 09/163,115  
; EARLIER FILING DATE:  
; NUMBER OF SEQ ID NOS: 15  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 2  
; LENGTH: 302  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-09-221-928-2

Query Match 17.7%; Score 443; DB 3; Length 302;  
Best Local Similarity 34.6%; Pred. No. 2.9e-27;  
Matches 90; Conservative 58; Mismatches 102; Indels 10; Gaps 6;  
QY 31 LOOKLGGSGFGTYLVSDKKAKGEBELKV-LKEISVGEEL-NPNETVQANLEAQLLSKLDH 88  
DB 36 IEKKIGRGQFSEVY---RAACLLDGVFVALKKVOIFDLMDAKARADCIKEIDLKQLNH 91  
QY 89 PAIVKPHASVFEQDNFCIITEYCEGRDLDDKIOEYKQAGKIFPENQIIEFWIQLLGV 148  
DB 92 PNVIKYASFIEDNELNIVLELADAGDLGRFFSSKTTAAHSLVGTPTVYMSPERIH 151  
QY 149 MHERRILHRDLKSKNVFL-KNNLLKIGDFGVSRLLMGSCDLATLTGTPTHYMSPEALKHQ 207  
DB 152 MHSRRVHRDIKPNANVITATGVVKGDLGLGRFFSSKTTAAHSLVGTPTVYMSPERIH 211  
QY 208 GYDTSKDIWSLACILYEMCCMHAFAGS--NFLSIVLKIIVEGDTSPSL-ERYPKELNALME 264  
DB 212 GYNFKSDIWSLGLLYEMAAALQSPFYGDKMNLKIEQCDYPPPLPSDHYSEELRQ 271  
QY 265 ESMLNKNPSLRPSAIELKI 284  
DB 272 NMCINPDEKRPDVTYVDV 291

[illegible]

us-10-803-278-4.rai

Mon Oct 4 10:35:28 2004

EARLIER APPLICATION NUMBER: US 09/163,115  
 EARLIER FILING DATE: 1998-09-29  
 NUMBER OF SEQ ID NOS: 15  
 SOFTWARE: Patent in Ver. 2.0  
 SEQ ID NO 2  
 LENGTH: 302  
 TYPE: PRT  
 ORGANISM: Homo sapiens  
 US-09-221-245-2

Query Match 17.7%; Score 443; DB 3; Length 302;  
 Best Local Similarity 34.6%; Pred. No. 2.9e-27;  
 Matches 90; Conservative 58; Mismatches 102; Indels 10; Gaps 6;  
 QY 31 LOOKLGSFGTGYLVSDKKAKGGEELKV-LKEISVGEI-NPNETVQANLEAQLSKLDH 88  
 Db 36 IEKKIGRGQSEVY----RAACLDGVPVALKKQVIFDLMDAKARADCKEIDLKQLNH 91  
 QY 89 PAIVKPHASFEQDNFCIIIEYCEGRDLDKIQEYKQAGKIPFENQIIIEWFIQILLGVYD 148  
 Db 92 PNVIKYASFIEDNELNIVLEADAGLSRMKHKFKQKRLIPERTVWKYFVQLCSALEH 151  
 QY 149 MHERRILHRLKSKNVFL-KNNLLKIGDFGVSRLLMGSCDLATTLTGTPHYMSPEALKHQ 207  
 Db 152 MHSRRVMHRDIKPNVFIATGVVKGDLGLGRFFSSKTTAAHSLVGTPTVYMSPERIHEN 211  
 QY 208 GYDTKSDIWSLACILYEMCMNHAFAGS--NFLSIVLKIVEGDTPSLP-ERYPKELNAIN 264  
 Db 212 GYNFKSDIWSLGLCLLYEMAALQSPFYGDKNLYSLCKIEQCDYPPPLSDHYSEELRLQV 271  
 QY 265 ESMLNKNPSLRPSAIEILKI 284  
 Db 272 NMCINPDPEKRPDVTYVYDV 291

RESULT 14  
 US-09-163-115-2  
 Sequence 2, Application US/09163115A  
 Patent No. 6183962  
 GENERAL INFORMATION:  
 APPLICANT: Acton, Susan  
 TITLE OF INVENTION: NOVEL CSAPK-1 NUCLEIC ACID MOLECULES AND USES THEREFOR  
 FILE REFERENCE: NMI-050  
 CURRENT APPLICATION NUMBER: US/09/163,115A  
 CURRENT FILING DATE: 1998-09-29  
 NUMBER OF SEQ ID NOS: 15  
 SOFTWARE: Patent in Ver. 2.0  
 SEQ ID NO 2  
 LENGTH: 302  
 TYPE: PRT  
 ORGANISM: Homo sapiens  
 US-09-163-115-2

Query Match 17.7%; Score 443; DB 3; Length 302;  
 Best Local Similarity 34.6%; Pred. No. 2.9e-27;  
 Matches 90; Conservative 58; Mismatches 102; Indels 10; Gaps 6;  
 QY 31 LOOKLGSFGTGYLVSDKKAKGGEELKV-LKEISVGEI-NPNETVQANLEAQLSKLDH 88  
 Db 36 IEKKIGRGQSEVY----RAACLDGVPVALKKQVIFDLMDAKARADCKEIDLKQLNH 91  
 QY 89 PAIVKPHASFEQDNFCIIIEYCEGRDLDKIQEYKQAGKIPFENQIIIEWFIQILLGVYD 148  
 Db 92 PNVIKYASFIEDNELNIVLEADAGLSRMKHKFKQKRLIPERTVWKYFVQLCSALEH 151  
 QY 149 MHERRILHRLKSKNVFL-KNNLLKIGDFGVSRLLMGSCDLATTLTGTPHYMSPEALKHQ 207  
 Db 152 MHSRRVMHRDIKPNVFIATGVVKGDLGLGRFFSSKTTAAHSLVGTPTVYMSPERIHEN 211  
 QY 208 GYDTKSDIWSLACILYEMCMNHAFAGS--NFLSIVLKIVEGDTPSLP-ERYPKELNAIN 264  
 Db 212 GYNFKSDIWSLGLCLLYEMAALQSPFYGDKNLYSLCKIEQCDYPPPLSDHYSEELRLQV 271

QY 265 ESMLNKNPSLRPSAIEILKI 284  
 Db 272 NMCINPDPEKRPDVTYVYDV 291  
 RESULT 15  
 US-09-221-528-2  
 Sequence 2, Application US/09221528  
 Patent No. 6190874  
 GENERAL INFORMATION:  
 APPLICANT: Acton, Susan  
 TITLE OF INVENTION: NOVEL CSAPK-1 NUCLEIC ACID MOLECULES AND USES THEREFOR  
 FILE REFERENCE: NMI-050  
 CURRENT APPLICATION NUMBER: US/09/221,528  
 CURRENT FILING DATE: 1998-12-28  
 EARLIER APPLICATION NUMBER: 09/163,115  
 EARLIER FILING DATE: 1998-09-29  
 NUMBER OF SEQ ID NOS: 15  
 SOFTWARE: Patent in Ver. 2.0  
 SEQ ID NO 2  
 LENGTH: 302  
 TYPE: PRT  
 ORGANISM: Homo sapiens  
 US-09-221-528-2

Query Match 17.7%; Score 443; DB 3; Length 302;  
 Best Local Similarity 34.6%; Pred. No. 2.9e-27;  
 Matches 90; Conservative 58; Mismatches 102; Indels 10; Gaps 6;  
 QY 31 LOOKLGSFGTGYLVSDKKAKGGEELKV-LKEISVGEI-NPNETVQANLEAQLSKLDH 88  
 Db 36 IEKKIGRGQSEVY----RAACLDGVPVALKKQVIFDLMDAKARADCKEIDLKQLNH 91  
 QY 89 PAIVKPHASFEQDNFCIIIEYCEGRDLDKIQEYKQAGKIPFENQIIIEWFIQILLGVYD 148  
 Db 92 PNVIKYASFIEDNELNIVLEADAGLSRMKHKFKQKRLIPERTVWKYFVQLCSALEH 151  
 QY 149 MHERRILHRLKSKNVFL-KNNLLKIGDFGVSRLLMGSCDLATTLTGTPHYMSPEALKHQ 207  
 Db 152 MHSRRVMHRDIKPNVFIATGVVKGDLGLGRFFSSKTTAAHSLVGTPTVYMSPERIHEN 211  
 QY 208 GYDTKSDIWSLACILYEMCMNHAFAGS--NFLSIVLKIVEGDTPSLP-ERYPKELNAIN 264  
 Db 212 GYNFKSDIWSLGLCLLYEMAALQSPFYGDKNLYSLCKIEQCDYPPPLSDHYSEELRLQV 271  
 QY 265 ESMLNKNPSLRPSAIEILKI 284  
 Db 272 NMCINPDPEKRPDVTYVYDV 291

Search completed: October 1, 2004, 11:22:09  
 Job time : 34 secs



GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 1, 2004, 11:20:51 ; Search time 135 Seconds  
(without alignments)  
1148.942 Million cell updates/sec

Title: US-10-803-278-4

Perfect score: 2501

Sequence: 1 MLKQEAARKVSGSGTAISTY.....LGYHGDGNLSLDEYWKNEK 482

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1351062 seqs, 321799191 residues

Total number of hits satisfying chosen parameters: 1351062

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

.Listing first 45 summaries

Database :

- Published Applications AA:\*  
1: /cgn2\_6/ptodata/2/pubpaa/US07\_PUBCOMB.pep.\*  
2: /cgn2\_6/ptodata/2/pubpaa/PCT\_NEW\_PUB.pep.\*  
3: /cgn2\_6/ptodata/2/pubpaa/US06\_NEW\_PUB.pep.\*  
4: /cgn2\_6/ptodata/2/pubpaa/US06\_PUBCOMB.pep.\*  
5: /cgn2\_6/ptodata/2/pubpaa/US07\_NEW\_PUB.pep.\*  
6: /cgn2\_6/ptodata/2/pubpaa/PCTUS\_PUBCOMB.pep.\*  
7: /cgn2\_6/ptodata/2/pubpaa/US08\_NEW\_PUB.pep.\*  
8: /cgn2\_6/ptodata/2/pubpaa/US08\_PUBCOMB.pep.\*  
9: /cgn2\_6/ptodata/2/pubpaa/US09A\_PUBCOMB.pep.\*  
10: /cgn2\_6/ptodata/2/pubpaa/US09B\_PUBCOMB.pep.\*  
11: /cgn2\_6/ptodata/2/pubpaa/US09C\_PUBCOMB.pep.\*  
12: /cgn2\_6/ptodata/2/pubpaa/US09\_NEW\_PUB.pep.\*  
13: /cgn2\_6/ptodata/2/pubpaa/US10A\_PUBCOMB.pep.\*  
14: /cgn2\_6/ptodata/2/pubpaa/US10B\_PUBCOMB.pep.\*  
15: /cgn2\_6/ptodata/2/pubpaa/US10C\_PUBCOMB.pep.\*  
16: /cgn2\_6/ptodata/2/pubpaa/US10\_NEW\_PUB.pep.\*  
17: /cgn2\_6/ptodata/2/pubpaa/US60\_NEW\_PUB.pep.\*  
18: /cgn2\_6/ptodata/2/pubpaa/US60\_PUBCOMB.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	2501	100.0	482	16	US-10-803-278-4
2	2408	96.3	645	12	US-10-333-314-5
3	2408	96.3	645	16	US-10-803-278-2
4	2408	96.3	653	12	US-10-425-114-37518
5	2408	96.3	653	12	US-10-425-114-53830
6	2405	96.2	645	9	US-09-731-231A-2
7	2405	96.2	645	16	US-10-751-985-2
8	1799.5	72.0	502	12	US-10-425-114-54112
9	1491	59.6	467	10	US-09-866-050A-513
10	1480	59.2	463	16	US-09-731-231A-4
11	1480	59.2	463	16	US-10-751-985-4
12	615	24.6	841	14	US-10-198-070-87
13	615	24.6	850	12	US-10-425-114-56926
14	615	24.6	850	12	US-10-425-114-56928
15	613	24.5	841	14	US-10-198-070-52

16	613	24.5	841	16	US-10-408-765A-1685	Sequence 1685, Ap
17	613	24.5	841	16	US-10-620-052A-32	Sequence 32, Appl
18	600.5	24.0	416	9	US-09-731-231A-6	Sequence 6, Appli
19	600.5	24.0	416	16	US-10-751-985-6	Sequence 6, Appli
20	598.5	23.9	489	12	US-10-114-270-182	Sequence 182, App
21	598	23.9	616	14	US-10-243-735-4	Sequence 4, Appli
22	598	23.9	616	16	US-10-730-010-4	Sequence 5, Appli
23	598	23.9	774	14	US-10-162-706-5	Sequence 5, Appli
24	593.5	23.7	260	9	US-09-731-231A-5	Sequence 5, Appli
25	593.5	23.7	260	16	US-10-751-985-5	Sequence 5, Appli
26	593.5	23.7	425	12	US-10-114-270-184	Sequence 184, App
27	592	23.7	506	12	US-10-114-270-180	Sequence 180, App
28	585.5	23.4	255	10	US-09-898-837A-36	Sequence 36, Appl
29	585.5	23.4	256	16	US-10-620-052A-48	Sequence 48, Appl
30	585.5	23.4	654	9	US-09-940-921B-4	Sequence 4, Appli
31	584	23.4	506	12	US-10-415-011-7	Sequence 7, Appli
32	584	23.4	640	14	US-10-243-735-2	Sequence 2, Appli
33	584	23.4	640	16	US-10-730-010-2	Sequence 2, Appli
34	584	23.4	683	9	US-09-940-921B-2	Sequence 2, Appli
35	582.5	23.3	1214	9	US-09-783-320-4	Sequence 4, Appli
36	575.5	23.0	403	12	US-10-114-270-188	Sequence 188, App
37	563.5	22.5	399	12	US-10-114-270-186	Sequence 186, App
38	559	22.4	254	10	US-09-898-837A-35	Sequence 35, Appl
39	539	21.6	459	9	US-09-771-161A-198	Sequence 198, App
40	533	21.3	107	11	US-09-764-875-658	Sequence 658, App
41	514	20.6	943	16	US-10-437-963-178155	Sequence 178155, A
42	509.5	20.4	584	12	US-10-425-114-64177	Sequence 64177, A
43	506.5	20.3	585	16	US-10-437-963-138823	Sequence 138823, A
44	505.5	20.2	321	12	US-10-424-599-274838	Sequence 274838, A
45	497	19.9	583	16	US-10-437-963-120489	Sequence 120489, A

ALIGNMENTS

RESULT 1

US-10-803-278-4  
; Sequence 4, Application US/10803278  
; Publication No. US20040175749A1  
; GENERAL INFORMATION:  
; APPLICANT: Walke, D. Wade  
; APPLICANT: Scoville, John  
; APPLICANT: Fiddle, Carl Johan  
; TITLE OF INVENTION: Novel Human Kinases and Polynucleotides Encoding the Same  
; FILE REFERENCE: LEX-0348-USA  
; CURRENT APPLICATION NUMBER: US/10/803,278  
; CURRENT FILING DATE: 2004-03-18  
; PRIOR APPLICATION NUMBER: US/10/196,927  
; FILING DATE: 2002-05-20  
; PRIOR APPLICATION NUMBER: US 60/293,248  
; PRIOR FILING DATE: 2001-05-24  
; NUMBER OF SEQ ID NOS: 5  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 4  
; LENGTH: 482  
; TYPE: PRT  
; ORGANISM: homo sapiens  
US-10-803-278-4

Query Match	100.0%;	Score	2501;	DB	16;	Length	482;
Best Local Similarity	100.0%;	Pred. No.	2.9e-163;				
Matches	482;	Conservative	0;	Mismatches	0;	Indels	0;
Gaps	0;						
Qy	1	MLKQEAARKVSGSGTAISTYPTKLIARRVYLQKLGSGSGFTVYVLSDDKAKRGEELKVL	60				
Db	1	MLKQEAARKVSGSGTAISTYPTKLIARRVYLQKLGSGSGFTVYVLSDDKAKRGEELKVL	60				
Qy	61	KEISVGLNPNETVQANLEAQLLSKLDHPAIVKHFASFVEQDNFCIITEYCEGRDLDKDI	120				
Db	61	KEISVGLNPNETVQANLEAQLLSKLDHPAIVKHFASFVEQDNFCIITEYCEGRDLDKDI	120				
Qy	121	QEKYQAKIPENCIIETFWIALLGVYMERILHDLKSNVFLKNLLKGDGVSR	180				
Db	121	QEKYQAKIPENCIIETFWIALLGVYMERILHDLKSNVFLKNLLKGDGVSR	180				

```
Db 121 QEQKQAGKIPFENQIIEWFIQLLGVYDVMHERRILHRDLKSKNVFLKNNLLKIGDFGVSR 180
Qy 181 LLMGSCDLATLTGTGTHYMSPEALKHQGYDTKSDIWSLACILYEMCMNHAFAGSNFLSI 240
Db 181 LLMGSCDLATLTGTGTHYMSPEALKHQGYDTKSDIWSLACILYEMCMNHAFAGSNFLSI 240
Qy 241 VLKIVEGDTPSLPERYPKELNAIMESMLNKNPSLRPSAIEILKIPYLDQQLNLMCRYSE 300
Db 241 VLKIVEGDTPSLPERYPKELNAIMESMLNKNPSLRPSAIEILKIPYLDQQLNLMCRYSE 300
Qy 301 MTLKEDKNDLCKQEAHIIINAMQKRIHLQTLRALSEVQKMTPRRMLRLKLAADKARKL 360
Db 301 MTLKEDKNDLCKQEAHIIINAMQKRIHLQTLRALSEVQKMTPRRMLRLKLAADKARKL 360
Qy 361 KKIIVEKYENSXRMQELSRNFQQLSDVVLHKKTHLKGMEKEKEQPEGRILSCSPQDEDE 420
Db 361 KKIIVEKYENSXRMQELSRNFQQLSDVVLHKKTHLKGMEKEKEQPEGRILSCSPQDEDE 420
Qy 421 ERWQGREESDEPTLENLPESQIPSPMDLHELESIVEDATSDLYGDCNLIISLDEYWN 480
Db 421 ERWQGREESDEPTLENLPESQIPSPMDLHELESIVEDATSDLYGDCNLIISLDEYWN 480
Qy 481 EK 482
Db 481 EK 482

RESULT 2
US-10-333-314-5
; Sequence 5, Application US/10333314
; Publication No. US20030211093A1
; GENERAL INFORMATION:
; APPLICANT: INCYTE GENOMICS, INC.; YUE, Henry
; APPLICANT: KHAN, Farrah A.; GURURAJAN, Rajagopal
; APPLICANT: HAFALIA, April J.A.; CHAWLA, Narinder K.
; APPLICANT: ARVIZU, Chandra S.; RAMKUMAR, Jayalaxmi
; APPLICANT: GANDHI, Ameena R.; POLICKY, Jennifer L.
; APPLICANT: BAUGHN, Mariah R.; TRIBOULEY, Catherine M.
; APPLICANT: THORNTON, Michael B.; BANDMAN, Olga
; APPLICANT: NGUYEN, Daniel B.; LU, Yan
; APPLICANT: BURFORD, Neil; LAL, Preeti G.
; APPLICANT: ELLIOTT, Vicki S.; RECIPON, Shirley A.
; APPLICANT: KEARNEY, Liam; LU, Dying Aina M.
; APPLICANT: GREENWALD, Sara R.; TANG, Y. Tom
; APPLICANT: XU, Yuming; WALSH, Roderick T.
; APPLICANT: GRETZEN, Kimberly J.; YANG, Junming
; APPLICANT: JACKSON, Jennifer L.
; TITLE OF INVENTION: HUMAN KINASES
; FILE REFERENCE: PI-0162 USN
; CURRENT APPLICATION NUMBER: US/10/333,314
; CURRENT FILING DATE: 2003-01-15
; PRIOR APPLICATION NUMBER: PCT/US01/23092
; PRIOR FILING DATE: 2001-07-20
; PRIOR APPLICATION NUMBER: US 60/220,038
; PRIOR FILING DATE: 2000-07-21
; PRIOR APPLICATION NUMBER: US 60/222,112
; PRIOR FILING DATE: 2000-07-28
; PRIOR APPLICATION NUMBER: US 60/222,831
; PRIOR FILING DATE: 2000-08-04
; PRIOR APPLICATION NUMBER: US 60/224,729
; PRIOR FILING DATE: 2000-08-11
; NUMBER OF SEQ ID NOS: 40
; SOFTWARE: PERL Program
; SEQ ID NO 5
; LENGTH: 645
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; OTHER INFORMATION: Incyte ID No: 7477585CD1
US-10-333-314-5
```

```
Query Match
Best Local Similarity 96.3%; Score 2408; DB 12; Length 645;
Matches 466; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 MLKFQEAACVSGSSTAISTYPTKTLIARRVYVLOQKLGSGSGFTVYVSDKKAKRGEELKVL 60
Db 1 MLKFQEAACVSGSSTAISTYPTKTLIARRVYVLOQKLGSGSGFTVYVSDKKAKRGEELKVL 60
Qy 61 KEISVGLNPNETVQANLEAQLLSKLDHPAIVKHFASFVEQDNFCIITEYCEGRDLDDKI 120
Db 61 KEISVGLNPNETVQANLEAQLLSKLDHPAIVKHFASFVEQDNFCIITEYCEGRDLDDKI 120
Qy 121 QEQKQAGKIPFENQIIEWFIQLLGVYDVMHERRILHRDLKSKNVFLKNNLLKIGDFGVSR 180
Db 121 QEQKQAGKIPFENQIIEWFIQLLGVYDVMHERRILHRDLKSKNVFLKNNLLKIGDFGVSR 180
Qy 181 LLMGSCDLATLTGTGTHYMSPEALKHQGYDTKSDIWSLACILYEMCMNHAFAGSNFLSI 240
Db 181 LLMGSCDLATLTGTGTHYMSPEALKHQGYDTKSDIWSLACILYEMCMNHAFAGSNFLSI 240
Qy 241 VLKIVEGDTPSLPERYPKELNAIMESMLNKNPSLRPSAIEILKIPYLDQQLNLMCRYSE 300
Db 241 VLKIVEGDTPSLPERYPKELNAIMESMLNKNPSLRPSAIEILKIPYLDQQLNLMCRYSE 300
Qy 301 MTLKEDKNDLCKQEAHIIINAMQKRIHLQTLRALSEVQKMTPRRMLRLKLAADKARKL 360
Db 301 MTLKEDKNDLCKQEAHIIINAMQKRIHLQTLRALSEVQKMTPRRMLRLKLAADKARKL 360
Qy 361 KKIIVEKYENSXRMQELSRNFQQLSDVVLHKKTHLKGMEKEKEQPEGRILSCSPQDEDE 420
Db 361 KKIIVEKYENSXRMQELSRNFQQLSDVVLHKKTHLKGMEKEKEQPEGRILSCSPQDEDE 420
Qy 421 ERWQGREESDEPTLENLPESQIPSPMDLHELESIVEDATSDLYG 466
Db 421 ERWQGREESDEPTLENLPESQIPSPMDLHELESIVEDATSDLYG 466

RESULT 3
US-10-803-278-2
; Sequence 2, Application US/10803278
; Publication No. US20040175749A1
; GENERAL INFORMATION:
; APPLICANT: Walke, D. Wade
; APPLICANT: Scoville, John
; APPLICANT: Friddle, Carl Johan
; TITLE OF INVENTION: Novel Human Kinases and Polynucleotides Encoding the Same
; FILE REFERENCE: LEX-0348-USA
; CURRENT APPLICATION NUMBER: US/10/803,278
; CURRENT FILING DATE: 2004-03-18
; PRIOR APPLICATION NUMBER: US/10/196,927
; PRIOR FILING DATE: 2002-05-20
; PRIOR APPLICATION NUMBER: US 60/293,248
; PRIOR FILING DATE: 2001-05-24
; NUMBER OF SEQ ID NOS: 5
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 645
; TYPE: PRT
; ORGANISM: homo sapiens
US-10-803-278-2

Query Match
Best Local Similarity 96.3%; Score 2408; DB 16; Length 645;
Matches 466; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 MLKFQEAACVSGSSTAISTYPTKTLIARRVYVLOQKLGSGSGFTVYVSDKKAKRGEELKVL 60
Db 1 MLKFQEAACVSGSSTAISTYPTKTLIARRVYVLOQKLGSGSGFTVYVSDKKAKRGEELKVL 60
Qy 61 KEISVGLNPNETVQANLEAQLLSKLDHPAIVKHFASFVEQDNFCIITEYCEGRDLDDKI 120
Db 61 KEISVGLNPNETVQANLEAQLLSKLDHPAIVKHFASFVEQDNFCIITEYCEGRDLDDKI 120
```

QY 121 QEYKQAGKIPFENQIIEWFIQLLGVYMHERRILHRDLKSNVFLKNLLKIGDFGYSR 180  
DB 121 QEYKQAGKIPFENQIIEWFIQLLGVYMHERRILHRDLKSNVFLKNLLKIGDFGYSR 180  
QY 181 LLMGSCDLATTTGTGTPHYNSPEALKHQGYDTKSDIWSLACILYEMCCMNHAFAGSNFLSI 240  
DB 181 LLMGSCDLATTTGTGTPHYNSPEALKHQGYDTKSDIWSLACILYEMCCMNHAFAGSNFLSI 240  
QY 241 VLKIVEGDTFSLPERYPKELNAMESMLNKNPSLRPSAIELIKIPYLBOLQNLKCRYSR 300  
DB 241 VLKIVEGDTFSLPERYPKELNAMESMLNKNPSLRPSAIELIKIPYLBOLQNLKCRYSR 300  
QY 301 MTEEDKNLDCQKEAAHIIINAMQKRIHLQTLRALSEVQKMPREMRRLKLOAADEKARKL 360  
DB 301 MTEEDKNLDCQKEAAHIIINAMQKRIHLQTLRALSEVQKMPREMRRLKLOAADEKARKL 360  
QY 361 KXIVEKYEENSKMQLSRNFQQLSVDLVHEKTHLKGMEKEEOPGRLSCSPQDEDE 420  
DB 361 KXIVEKYEENSKMQLSRNFQQLSVDLVHEKTHLKGMEKEEOPGRLSCSPQDEDE 420  
QY 421 ERWQGREESDEPTLENLPESQIPSPMDLHELESIVEDATSDLGYN 466  
DB 421 ERWQGREESDEPTLENLPESQIPSPMDLHELESIVEDATSDLGYN 466

RESULT 4  
US-10-425-114-37518  
; Sequence 37518, Application US/10425114  
; Publication No. US20040034888A1  
; GENERAL INFORMATION:  
; APPLICANT: Liu, Jingdong  
; APPLICANT: Zhou, Yihua  
; APPLICANT: Kovalic, David K.  
; APPLICANT: Screen, Steven E  
; APPLICANT: Tabaska, Jack E  
; APPLICANT: Cao, Yongwei  
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With  
; FILE REFERENCE: 38-21(53313)B  
; CURRENT APPLICATION NUMBER: US/10/425,114  
; CURRENT FILING DATE: 2003-04-28  
; NUMBER OF SEQ ID NOS: 73128  
; SEQ ID NO 37518  
; LENGTH: 653  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
; FEATURE:  
; OTHER INFORMATION: Clone ID: LIB3547-006-G1\_FLI.pep  
US-10-425-114-37518

Query Match 96.3%; Score 2408; DB 12; Length 653;  
Best Local Similarity 100.0%; Pred. No. 1e-156;  
Matches 466; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MLKQEAACVSGSSTAISTYPTKLIARRVYVLOQLGSGSGFTVYVSDKAKRGEELKVL 60  
DB 9 MLKQEAACVSGSSTAISTYPTKLIARRVYVLOQLGSGSGFTVYVSDKAKRGEELKVL 68  
QY 61 KEISVGEINPNETVQANLEAQLSKLDHPAIVKPHASFVEQDNFCIITEYCEGRDLDDKI 120  
DB 69 KEISVGEINPNETVQANLEAQLSKLDHPAIVKPHASFVEQDNFCIITEYCEGRDLDDKI 128  
QY 121 QEYKQAGKIPFENQIIEWFIQLLGVYMHERRILHRDLKSNVFLKNLLKIGDFGYSR 180  
DB 129 QEYKQAGKIPFENQIIEWFIQLLGVYMHERRILHRDLKSNVFLKNLLKIGDFGYSR 188  
QY 181 LLMGSCDLATTTGTGTPHYNSPEALKHQGYDTKSDIWSLACILYEMCCMNHAFAGSNFLSI 240  
DB 189 LLMGSCDLATTTGTGTPHYNSPEALKHQGYDTKSDIWSLACILYEMCCMNHAFAGSNFLSI 248  
QY 241 VLKIVEGDTFSLPERYPKELNAMESMLNKNPSLRPSAIELIKIPYLBOLQNLKCRYSR 300  
DB 249 VLKIVEGDTFSLPERYPKELNAMESMLNKNPSLRPSAIELIKIPYLBOLQNLKCRYSR 308  
QY 301 MTEEDKNLDCQKEAAHIIINAMQKRIHLQTLRALSEVQKMPREMRRLKLOAADEKARKL 360  
DB 309 QEYKQAGKIPFENQIIEWFIQLLGVYMHERRILHRDLKSNVFLKNLLKIGDFGYSR 368  
QY 361 KXIVEKYEENSKMQLSRNFQQLSVDLVHEKTHLKGMEKEEOPGRLSCSPQDEDE 420  
DB 249 VLKIVEGDTFSLPERYPKELNAMESMLNKNPSLRPSAIELIKIPYLBOLQNLKCRYSR 308

QY 301 MTEEDKNLDCQKEAAHIIINAMQKRIHLQTLRALSEVQKMPREMRRLKLOAADEKARKL 360  
DB 309 MTEEDKNLDCQKEAAHIIINAMQKRIHLQTLRALSEVQKMPREMRRLKLOAADEKARKL 368  
QY 361 KXIVEKYEENSKMQLSRNFQQLSVDLVHEKTHLKGMEKEEOPGRLSCSPQDEDE 420  
DB 369 KXIVEKYEENSKMQLSRNFQQLSVDLVHEKTHLKGMEKEEOPGRLSCSPQDEDE 428  
QY 421 ERWQGREESDEPTLENLPESQIPSPMDLHELESIVEDATSDLGYN 466  
DB 429 ERWQGREESDEPTLENLPESQIPSPMDLHELESIVEDATSDLGYN 474

RESULT 5  
US-10-425-114-53830  
; Sequence 53830, Application US/10425114  
; Publication No. US20040034888A1  
; GENERAL INFORMATION:  
; APPLICANT: Liu, Jingdong  
; APPLICANT: Zhou, Yihua  
; APPLICANT: Kovalic, David K.  
; APPLICANT: Screen, Steven E  
; APPLICANT: Tabaska, Jack E  
; APPLICANT: Cao, Yongwei  
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With  
; FILE REFERENCE: 38-21(53313)B  
; CURRENT APPLICATION NUMBER: US/10/425,114  
; CURRENT FILING DATE: 2003-04-28  
; NUMBER OF SEQ ID NOS: 73128  
; SEQ ID NO 53830  
; LENGTH: 653  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
; FEATURE:  
; OTHER INFORMATION: Clone ID: LIB3119-029-B6\_FLI.pep  
US-10-425-114-53830

Query Match 96.3%; Score 2408; DB 12; Length 653;  
Best Local Similarity 100.0%; Pred. No. 1e-156;  
Matches 466; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MLKQEAACVSGSSTAISTYPTKLIARRVYVLOQLGSGSGFTVYVSDKAKRGEELKVL 60  
DB 9 MLKQEAACVSGSSTAISTYPTKLIARRVYVLOQLGSGSGFTVYVSDKAKRGEELKVL 68  
QY 61 KEISVGEINPNETVQANLEAQLSKLDHPAIVKPHASFVEQDNFCIITEYCEGRDLDDKI 120  
DB 69 KEISVGEINPNETVQANLEAQLSKLDHPAIVKPHASFVEQDNFCIITEYCEGRDLDDKI 128  
QY 121 QEYKQAGKIPFENQIIEWFIQLLGVYMHERRILHRDLKSNVFLKNLLKIGDFGYSR 180  
DB 129 QEYKQAGKIPFENQIIEWFIQLLGVYMHERRILHRDLKSNVFLKNLLKIGDFGYSR 188  
QY 181 LLMGSCDLATTTGTGTPHYNSPEALKHQGYDTKSDIWSLACILYEMCCMNHAFAGSNFLSI 240  
DB 189 LLMGSCDLATTTGTGTPHYNSPEALKHQGYDTKSDIWSLACILYEMCCMNHAFAGSNFLSI 248  
QY 241 VLKIVEGDTFSLPERYPKELNAMESMLNKNPSLRPSAIELIKIPYLBOLQNLKCRYSR 300  
DB 249 VLKIVEGDTFSLPERYPKELNAMESMLNKNPSLRPSAIELIKIPYLBOLQNLKCRYSR 308  
QY 301 MTEEDKNLDCQKEAAHIIINAMQKRIHLQTLRALSEVQKMPREMRRLKLOAADEKARKL 360  
DB 309 MTEEDKNLDCQKEAAHIIINAMQKRIHLQTLRALSEVQKMPREMRRLKLOAADEKARKL 368  
QY 361 KXIVEKYEENSKMQLSRNFQQLSVDLVHEKTHLKGMEKEEOPGRLSCSPQDEDE 420  
DB 369 KXIVEKYEENSKMQLSRNFQQLSVDLVHEKTHLKGMEKEEOPGRLSCSPQDEDE 428  
QY 421 ERWQGREESDEPTLENLPESQIPSPMDLHELESIVEDATSDLGYN 466

```
Db 429 ERWQGREESDEPTLENLPESQIPSMDLHELESIVEDATSDLGX 474
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 645
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-751-985-2

Query Match 96.2%; Score 2405; DB 16; Length 645;
Best Local Similarity 99.8%; Pred. No. 1.6e-156;
Matches 465; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 MLKFOBAACVSGSTAISTYPTKLIARRVYLQKLGSGSGFTVYLVSDDKAKRGSELKVL 60
DB 1 MLKFOBAACVSGSTAISTYPTKLIARRVYLQKLGSGSGFTVYLVSDDKAKRGSELKVL 60
QY 61 KEISVGLNPNETVQANLEAQLSKLDHPAIVKFXHASFVEQDNFCIITEYCEGRDLDDKI 120
DB 61 KEISVGLNPNETVQANLEAQLSKLDHPAIVKFXHASFVEQDNFCIITEYCEGRDLDDKI 120
QY 121 QEYKQAGKIPFENQIIEWFIQLLGVDMYHERRILHRDLKSNVFLKNLLKIGDFGVSR 180
DB 121 QEYKQAGKIPFENQIIEWFIQLLGVDMYHERRILHRDLKSNVFLKNLLKIGDFGVSR 180
QY 181 LMGSCDLATTLTGTHYMSPEALKHQGYDTKSDIWSLACILYEMCMNHAFAGSNFLSI 240
DB 181 LMGSCDLATTLTGTHYMSPEALKHQGYDTKSDIWSLACILYEMCMNHAFAGSNFLSI 240
QY 241 VLKIVGDTPLSPERYPKELNAMESMLKNPSLRPSAIEILKIPYLDOLQNLKRYSE 300
DB 241 VLKIVGDTPLSPERYPKELNAMESMLKNPSLRPSAIEILKIPYLDOLQNLKRYSE 300
QY 301 MTLEDKNLDCQKAAHIINAMQKRIHLQTLRALSEVQKMTPRERMLRKLQAADEKARKL 360
DB 301 MTLEDKNLDCQKAAHIINAMQKRIHLQTLRALSEVQKMTPRERMLRKLQAADEKARKL 360
QY 361 KKIIVEKYENS KRMQELSRNFQQLSVDLVHEKTHLKGMEKEQEPGRSLSCPODEDE 420
DB 361 KKIIVEKYENS KRMQELSRNFQQLSVDLVHEKTHLKGMEKEQEPGRSLSCPODEDE 420
QY 421 ERWQGREESDEPTLENLPESQIPSMDLHELESIVEDATSDLGX 466
DB 421 ERWQGREESDEPTLENLPESQIPSMDLHELESIVEDATSDLGX 466

RESULT 8
US-10-425-114-54112
; Sequence 54112, Application US/10425114
; Publication No. US20040034888A1
; GENERAL INFORMATION:
; APPLICANT: Zhou, Yihua
; APPLICANT: Kovalic, David K.
; APPLICANT: Screen, Steven E
; APPLICANT: Tabaska, Jack E
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53313)B
; CURRENT APPLICATION NUMBER: US/10/425,114
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 73128
; SEQ ID NO 54112
; LENGTH: 502
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: Clone ID: LIB3063-070-A3_FLI.pgp
US-10-425-114-54112

Query Match 72.0%; Score 1799.5; DB 12; Length 502;
Best Local Similarity 77.5%; Pred. No. 3.8e-115;
Matches 361; Conservative 0; Mismatches 0; Indels 105; Gaps 1;

Db 429 ERWQGREESDEPTLENLPESQIPSMDLHELESIVEDATSDLGX 474
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 645
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-751-985-2

Query Match 96.2%; Score 2405; DB 9; Length 645;
Best Local Similarity 99.8%; Pred. No. 1.6e-156;
Matches 465; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 MLKFOBAACVSGSTAISTYPTKLIARRVYLQKLGSGSGFTVYLVSDDKAKRGSELKVL 60
DB 1 MLKFOBAACVSGSTAISTYPTKLIARRVYLQKLGSGSGFTVYLVSDDKAKRGSELKVL 60
QY 61 KEISVGLNPNETVQANLEAQLSKLDHPAIVKFXHASFVEQDNFCIITEYCEGRDLDDKI 120
DB 61 KEISVGLNPNETVQANLEAQLSKLDHPAIVKFXHASFVEQDNFCIITEYCEGRDLDDKI 120
QY 121 QEYKQAGKIPFENQIIEWFIQLLGVDMYHERRILHRDLKSNVFLKNLLKIGDFGVSR 180
DB 121 QEYKQAGKIPFENQIIEWFIQLLGVDMYHERRILHRDLKSNVFLKNLLKIGDFGVSR 180
QY 181 LMGSCDLATTLTGTHYMSPEALKHQGYDTKSDIWSLACILYEMCMNHAFAGSNFLSI 240
DB 181 LMGSCDLATTLTGTHYMSPEALKHQGYDTKSDIWSLACILYEMCMNHAFAGSNFLSI 240
QY 241 VLKIVGDTPLSPERYPKELNAMESMLKNPSLRPSAIEILKIPYLDOLQNLKRYSE 300
DB 241 VLKIVGDTPLSPERYPKELNAMESMLKNPSLRPSAIEILKIPYLDOLQNLKRYSE 300
QY 301 MTLEDKNLDCQKAAHIINAMQKRIHLQTLRALSEVQKMTPRERMLRKLQAADEKARKL 360
DB 301 MTLEDKNLDCQKAAHIINAMQKRIHLQTLRALSEVQKMTPRERMLRKLQAADEKARKL 360
QY 361 KKIIVEKYENS KRMQELSRNFQQLSVDLVHEKTHLKGMEKEQEPGRSLSCPODEDE 420
DB 361 KKIIVEKYENS KRMQELSRNFQQLSVDLVHEKTHLKGMEKEQEPGRSLSCPODEDE 420
QY 421 ERWQGREESDEPTLENLPESQIPSMDLHELESIVEDATSDLGX 466
DB 421 ERWQGREESDEPTLENLPESQIPSMDLHELESIVEDATSDLGX 466

RESULT 7
US-10-751-985-2
; Sequence 2, Application US/10751985
; Publication No. US20040126861A1
; GENERAL INFORMATION:
; APPLICANT: GUGLER, Karl et al.
; TITLE OF INVENTION: ISOLATED HUMAN KINASE PROTEINS, NUCLEIC
; TITLE OF INVENTION: ACID MOLECULES ENCODING HUMAN KINASE PROTEINS, AND USES
; TITLE OF INVENTION: THEREOF
; FILE REFERENCE: CL001007C
; CURRENT APPLICATION NUMBER: US/10/751,985
; CURRENT FILING DATE: 2004-01-07
; NUMBER OF SEQ ID NOS: 6
```

QY 1 MLKFOBAKCVSGSTAISTAYPTKTLIARRVYVLOOKLGGSGFGTVYVSDKAKRGSELKYL 60  
DB 9 MLKFOBAKCVSGSTAISTAYPTKTLIARRVYVLOOKLGGSGFGTVYVSDKAKRGSELKYL 68  
QY 61 KEISVCELNPNETVQANLEAQLSLKLDHPAIVKFEHASFVEQDNFCIITEYCEGRDLDDKI 120  
DB 69 KEISVCELNPNETVQANLEAQLSLKLDHPAIVKFEHASFVEQDNFCIITEYCEGRDLDDKI 128  
QY 121 QYKQAGKIFPNQIIEWFIQLLGVYDVMHERRILHRDLKSNVFLKNNLLKIGDFGVSR 180  
DB 129 QYKQAGKIFPNQIIEWFIQLLGVYDVMHERRILHRDLKSNVFLKNNLLKIGDFGVSR 188  
QY 181 LLMGSCDLATTLTGTPHYMSPEALKHOGYDTKSDIWSLACILYEMCCMNHAFAGSNFLSI 240  
DB 189 LLMGSCDLATTLTGTPHYMSPEALKHOGYDTKSDIWSLACILYEMCCMNHAFAGSNFLSI 244  
QY 241 VLKIVEGDTSPSLPERYPKELNAINMESMLNKNPSPALRPSAIEILKIPYLDLQNLQNMCRYS 300  
DB 225 -----QELDELTLLESLSQPQLPCNLNDELEPSLEGTIVDLGHY 390  
QY 301 MTELEKXNLDCKEAAHIINAMOKRIHLQTLRALSEVQKMTPRERMLRLKLOAADEKARKL 360  
DB 225 -----QKRIHLQTLRALSEVQKMTPRERMLRLKLOAADEKARKL 263  
QY 361 KXIVEKYBENSKRMQELSRNFQOOLSVDVLHEKTHLKGMEKEKEQPEGRLLSCSPQDEDE 420  
DB 264 KXIVEKYBENSKRMQELSRNFQOOLSVDVLHEKTHLKGMEKEKEQPEGRLLSCSPQDEDE 323  
QY 421 ERWQGREESDPTLENLPESQIPSMDLHELESIVEDATSDLG 466  
DB 324 ERWQGREESDPTLENLPESQIPSMDLHELESIVEDATSDLG 369  
RESULT 9  
US-09-866-050A-513  
; Sequence 513, Application US/09866050A  
; Publication No. US2003004071A1  
; GENERAL INFORMATION:  
; APPLICANT: Watson, James D.  
; APPLICANT: Strachan, Lorna  
; APPLICANT: Sleeman, Matthew  
; APPLICANT: Orlust, Rene  
; APPLICANT: Murison, James G.  
; APPLICANT: Kumble, Krishanand D.  
; TITLE OF INVENTION: Compositions Isolated From Skin Cells  
; FILE REFERENCE: 11000.1011c4U  
; CURRENT APPLICATION NUMBER: US/09/866,050A  
; CURRENT FILING DATE: 2001-05-24  
; NUMBER OF SEQ ID NOS: 725  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 513  
; LENGTH: 467  
; TYPE: PRT  
; ORGANISM: Rat  
US-09-866-050A-513

Query Match 59.6%; Score 1491; DB 10; Length 467;  
Best Local Similarity 64.3%; Pred. No. 5e-94;  
Matches 301; Conservative 40; Mismatches 47; Indels 80; Gaps 5;

QY 1 MLKFOBAKCVSGSTAISTAYPTKTLIARRVYVLOOKLGGSGFGTVYVSDKAKRGSELKYL 60  
DB 1 MLKFOBAKCVSGSTAISTAYPTKTLIARRVYVLOOKLGGSGFGTVYVSDKAKRGSELKYL 58  
QY 61 KEISVCELNPNETVQANLEAQLSLKLDHPAIVKFEHASFVEQDNFCIITEYCEGRDLDDKI 120  
DB 59 KEISVCELNPNETVQANLEAQLSLKLDHPAIVKFEHASFVEQDNFCIITEYCEGRDLDDKI 118  
QY 121 QYKQAGKIFPNQIIEWFIQLLGVYDVMHERRILHRDLKSNVFLKNNLLKIGDFGVSR 180  
DB 119 QYKQAGKIFPNQIIEWFIQLLGVYDVMHERRILHRDLKSNVFLKNNLLKIGDFGVSR 178

QY 181 LLMGSCDLATTLTGTPHYMSPEALKHOGYDTKSDIWSLACILYEMCCMNHAFAGSNFLSI 240  
DB 179 LLMGSCDLATTLTGTPHYMSPEALKHOGYDTKSDIWSLACILYEMCCMNHAFAGSNFLSV 238  
QY 241 VLKIVEGDTSPSLPERYPKELNAINMESMLNKNPSPALRPSAIEILKIPYLDLQNLQNMCRYS 300  
DB 239 VLKIVEGDTSPSLPERYPKELNAINMESMLNKNPSPALRPSAIEILKIPYLDLQNLQNMCRYS 298  
QY 301 MTELEKXNLDCKEAAHIINAMOKRIHLQTLRALSEVQKMTPRERMLRLKLOAADEKARK 359  
DB 299 MTELEKXNLDCKEAAHIINAMOKRIHLQTLRALSEVQKMTPRERMLRLKLOAADEKARK 320  
QY 360 LKXIVEKYBENSKRMQELSRNFQOOLSVDVLHEKTHLKGMEKEKEQPEGRLLSCSPQDE 418  
DB 321 --KITEGRRYKNNKRAKELSRNFQOOLSVDVLHEKTHLKGMEKEKEQPEGRLLSCSPQDE 351  
QY 419 DEERWQGREESDPTLENLPESQIPSMDLHELESIVEDATSDLG 466  
DB 352 -----QELDELTLLESLSQPQLPCNLNDELEPSLEGTIVDLGHY 390

## RESULT 10

US-09-731-231A-4  
; Sequence 4, Application US/09731231A  
; Patent No. US20020082189A1  
; GENERAL INFORMATION:  
; APPLICANT: GUEGLER, Karl et al  
; TITLE OF INVENTION: ISOLATED HUMAN KINASE PROTEINS, NUCLEIC  
; TITLE OF INVENTION: ACID MOLECULES ENCODING HUMAN KINASE PROTEINS, AND USES  
; FILE REFERENCE: CL001007  
; CURRENT APPLICATION NUMBER: US/09/731,231A  
; CURRENT FILING DATE: 2000-12-07  
; NUMBER OF SEQ ID NOS: 6  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 4  
; LENGTH: 463  
; TYPE: PRT  
; ORGANISM: Human  
US-09-731-231A-4

Query Match 59.2%; Score 1480; DB 9; Length 463;  
Best Local Similarity 100.0%; Pred. No. 2.8e-93;  
Matches 284; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 183 MGSCLATTLTGTPHYMSPEALKHOGYDTKSDIWSLACILYEMCCMNHAFAGSNFLSV 242  
DB 1 MGSCLATTLTGTPHYMSPEALKHOGYDTKSDIWSLACILYEMCCMNHAFAGSNFLSV 60  
QY 243 KIVEGDTSPSLPERYPKELNAINMESMLNKNPSPALRPSAIEILKIPYLDLQNLQNMCRYS 302  
DB 61 KIVEGDTSPSLPERYPKELNAINMESMLNKNPSPALRPSAIEILKIPYLDLQNLQNMCRYS 120  
QY 303 LEDXNLDCKEAAHIINAMOKRIHLQTLRALSEVQKMTPRERMLRLKLOAADEKARKL 362  
DB 121 LEDXNLDCKEAAHIINAMOKRIHLQTLRALSEVQKMTPRERMLRLKLOAADEKARKL 180  
QY 363 IVEKYBENSKRMQELSRNFQOOLSVDVLHEKTHLKGMEKEKEQPEGRLLSCSPQDEDE 422  
DB 181 IVEKYBENSKRMQELSRNFQOOLSVDVLHEKTHLKGMEKEKEQPEGRLLSCSPQDEDE 240  
QY 423 WQGREESDPTLENLPESQIPSMDLHELESIVEDATSDLG 466  
DB 241 WQGREESDPTLENLPESQIPSMDLHELESIVEDATSDLG 284

## RESULT 11

US-10-751-985-4  
; Sequence 4, Application US/10751985  
; Publication No. US20040126861A1  
; GENERAL INFORMATION:  
; APPLICANT: GUEGLER, Karl et al.  
; TITLE OF INVENTION: ISOLATED HUMAN KINASE PROTEINS, NUCLEIC

;; TITLE OF INVENTION: ACID MOLECULES ENCODING HUMAN KINASE PROTEINS, AND USES  
;; TITLE OF INVENTION: THEREOF  
;; FILE REFERENCE: CL001007CON  
;; CURRENT APPLICATION NUMBER: US/10/751,985  
;; CURRENT FILING DATE: 2004-01-07  
;; NUMBER OF SEQ ID NOS: 6  
;; SOFTWARE: FastSeq for Windows Version 4.0  
;; SEQ ID NO 4  
;; LENGTH: 463  
;; TYPE: PRT  
;; ORGANISM: Homo sapiens  
US-10-751-985-4

Query Match 59.2%; Score 1480; DB 16; Length 463;  
Best Local Similarity 100.0%; Pred. No. 2.8e-93;  
Matches 284; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 183 MGSCDLATLTGTPHYMSPEALKHQGYDTKSDIWSLACILYEMCCNNHAFAGSNFLSIVL 242  
DB 1 MGSCDLATLTGTPHYMSPEALKHQGYDTKSDIWSLACILYEMCCNNHAFAGSNFLSIVL 60

QY 243 KIVEGDTSPSPERYKELNAINESMLNKNPSLRPSAIEILKIPYLDQQLNLMCRYSEMT 302  
DB 61 KIVEGDTSPSPERYKELNAINESMLNKNPSLRPSAIEILKIPYLDQQLNLMCRYSEMT 120

QY 303 LEDKNLDCKOKEAAHINAKRIHLOTLRALSEVQKWTFRMRRLKQAADEKARKLK 362  
DB 121 LEDKNLDCKOKEAAHINAKRIHLOTLRALSEVQKWTFRMRRLKQAADEKARKLK 180

QY 363 IVEEKEYSKRMQELRSNFQOLSVDVLHKTGLKMGEEKEQEGRLSCSPQDEDER 422  
DB 181 IVEEKEYSKRMQELRSNFQOLSVDVLHKTGLKMGEEKEQEGRLSCSPQDEDER 240

QY 423 WQREESDEPTLENLPESQIPSDMLHELESIVEDATSDLG 466  
DB 241 WQREESDEPTLENLPESQIPSDMLHELESIVEDATSDLG 284

## RESULT 12

US-10-198-070-87  
; Sequence 87, Application US/10198070  
; Publication No. US20030109437A1  
; GENERAL INFORMATION:  
; APPLICANT: AVERBACK, PAUL  
; APPLICANT: GEMMELL, JACK  
; TITLE OF INVENTION: PEPTIDES EFFECTIVE IN THE TREATMENT OF TUMORS AND OTHER  
; TITLE OF INVENTION: CONDITIONS REQUIRING THE REMOVAL OR DESTRUCTION OF  
; TITLE OF INVENTION: CELLS  
; FILE REFERENCE: 59003.000008  
; CURRENT APPLICATION NUMBER: US/10/198,070  
; CURRENT FILING DATE: 2002-07-19  
; PRIOR APPLICATION NUMBER: 60/306,161  
; PRIOR FILING DATE: 2001-07-19  
; PRIOR APPLICATION NUMBER: 60/306,150  
; PRIOR FILING DATE: 2001-07-19  
; PRIOR APPLICATION NUMBER: 60/331,477  
; PRIOR FILING DATE: 2001-11-16  
; NUMBER OF SEQ ID NOS: 125  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 87  
; LENGTH: 841  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-10-198-070-87

Query Match 24.6%; Score 615; DB 14; Length 841;  
Best Local Similarity 34.7%; Pred. No. 1.1e-33;  
Matches 157; Conservative 87; Mismatches 152; Indels 56; Gaps 17;

QY 29 YVLOQKLGSGSFGTVYLVSDKKARGBELKVLKXISVGEINPNETVQANLEAQLSKLDH 88  
DB 6 YCYLRVVGKSGSYGEVTLV---KHRRDGKQYVIKKLNLRNASSRRRAAEQAEQLLSQLKH 62

QY 89 PAIVKPHASFVEODNFC-IITEYCEGRDLDDKIQEYKQAGKIPENQIIIEWFIQLLGLVD 147  
DB 63 PNIVTYKESWEGGDLIYVWVCEGDLRYRKE--QKQQLLPENQVVFQIWMALQ 120

QY 148 YMHERRILHRLDKSKNVFL-KNNLLKIGDGVSRLLMGSCDLATLTGTPHYMSPEALKH 206  
DB 121 YLHEKHILHRLDKTQNVFLTRTNIIVKVDLGIARVLENHCDMASTLIGTTPYMSPELFSN 180

QY 207 QGYDTKSDIWSLACILYEMCCNNHAFAGSNFLSIVLKIIVEGDTSPSPERYKELNAINES 266  
DB 181 KPNYKSDVWALGCCVYEMATLKHAFAKDNMSLVYRIIEGKLPAMPDRDYSPELAELIRT 240

QY 267 MLNKNPSLRPSAIEILKIPYLDQQLNLMCRYSEMTLED--KNLDCQ-KEAAHIN---- 319  
DB 241 MLSKRPERSVRSILRQPIKQISFFLEBATKITSKNNIKNGDSQSKPFAVWSGEAE 300

QY 320 AMQRIHLOTIRA-----LSEVQKWTFRMRRLKJ--QAADEKARKLKIVBEKYEN 371  
DB 301 SNHEVIHPQLSSEGSQTYIMGEGKCLSQEKPRASGLKSPASKAHTCKQDL-----SN 355

QY 372 SKRMQELRSNFQOLSVDVLHKTGLKMGEE-----KBEQPE-----GRLSCSPQD 417  
DB 356 ITELATISSVN-----IDL-----PAKGRDSVSDGVQENQPRYLDASNELGSI-CSISQ 405

QY 418 EDERWQGREESDEPTLENLPESQIPSDML 449  
DB 406 VEESMLQDNTKSSAQP--ENL---IPMWSSDI 432

## RESULT 13

US-10-425-114-56926  
; Sequence 56926, Application US/10425114  
; Publication No. US20040034888A1  
; GENERAL INFORMATION:  
; APPLICANT: Liu, Jingdong  
; APPLICANT: Zhou, Yihua  
; APPLICANT: Kovacic, David K.  
; APPLICANT: Screen, Steven E.  
; APPLICANT: Tabaska, Jack E.  
; APPLICANT: Cao, Yongwei  
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With  
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement  
; FILE REFERENCE: 38-21(53313)B  
; CURRENT APPLICATION NUMBER: US/10/425,114  
; CURRENT FILING DATE: 2003-04-28  
; NUMBER OF SEQ ID NOS: 73128  
; SEQ ID NO 56926  
; LENGTH: 850  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
; FEATURES:  
; OTHER INFORMATION: Clone ID: LIB4119-128-E10\_FLI pep  
US-10-425-114-56926

Query Match 24.6%; Score 615; DB 12; Length 850;  
Best Local Similarity 34.7%; Pred. No. 1.2e-33;  
Matches 157; Conservative 87; Mismatches 152; Indels 56; Gaps 17;

QY 29 YVLOQKLGSGSFGTVYLVSDKKARGBELKVLKXISVGEINPNETVQANLEAQLSKLDH 88  
DB 61 YCYLRVVGKSGSYGEVTLV---KHRRDGKQYVIKKLNLRNASSRRRAAEQAEQLLSQLKH 117

QY 89 PAIVKPHASFVEODNFC-IITEYCEGRDLDDKIQEYKQAGKIPENQIIIEWFIQLLGLVD 147  
DB 118 PNIVTYKESWEGGDLIYVWVCEGDLRYRKE--QKQQLLPENQVVFQIWMALQ 175

QY 148 YMHERRILHRLDKSKNVFL-KNNLLKIGDGVSRLLMGSCDLATLTGTPHYMSPEALKH 206  
DB 176 YLHEKHILHRLDKTQNVFLTRTNIIVKVDLGIARVLENHCDMASTLIGTTPYMSPELFSN 235

QY 207 QGYDTKSDIWSLACILYEMCCNNHAFAGSNFLSIVLKIIVEGDTSPSPERYKELNAINES 266  
DB 236 KPNYKSDVWALGCCVYEMATLKHAFAKDNMSLVYRIIEGKLPAMPDRDYSPELAELIRT 295





us-10-803-278-4.rapb

Mon Oct 4 10:35:29 2004

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: October 1, 2004, 11:05:15 ; Search time 25 Seconds  
(without alignments)  
1003.913 Million cell updates/sec

Title: US-10-803-278-4  
Perfect score: 2501  
Sequence: 1 MLKQEAQKCVSGSTAISTY.....LGVHGDCLISLDEYWKNEK 482

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 141681 seqs, 52070155 residues

Total number of hits satisfying chosen parameters: 141681

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : SwissProt\_42:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	615	24.6	841	1 NEK4_HUMAN	P51957 homo sapien
2	600.5	24.0	792	1 NEK4_MOUSE	Q921j2 mus musculus
3	598	23.9	774	1 NEK1_MOUSE	P51954 mus musculus
4	585	23.4	506	1 NEK3_HUMAN	P51956 homo sapien
5	584.5	23.3	511	1 NEK3_MOUSE	Q960a5 mus musculus
6	582.5	23.3	1258	1 NEK1_HUMAN	Q96py6 homo sapien
7	495	19.8	445	1 NEK2_HUMAN	P51955 homo sapien
8	474.5	19.0	697	1 NEK8_BRARE	P22209 brachydanio
9	463.5	18.5	435	1 KIX3_YEAST	P22209 saccharomyc
10	454	18.2	692	1 NEK8_HUMAN	Q86596 homo sapien
11	454	18.2	698	1 NEK8_MOUSE	Q912r4 mus musculus
12	452	18.1	443	1 NEK2_MOUSE	O35942 mus musculus
13	448	17.9	722	1 FIN1_SCHPO	O13839 schizosacch
14	445	17.8	302	1 NEK7_MOUSE	Q8es74 mus musculus
15	443	17.7	302	1 NEK7_HUMAN	Q8tdx7 homo sapien
16	443	17.7	313	1 NEK6_MOUSE	Q8es70 mus musculus
17	442	17.7	313	1 NEK6_RAT	P59895 rattus norv
18	440.5	17.6	699	1 NIMA_EMENT	P11837 emeritella
19	439	17.6	313	1 NEK6_HUMAN	Q9hc98 homo sapien
20	429.5	17.2	979	1 NEK9_HUMAN	Q8td19 homo sapien
21	425.5	17.0	984	1 NEK9_MOUSE	Q8klr7 mus musculus
22	414.5	16.6	944	1 NEK9_XENLA	Q7zcc8 xenopus lae
23	410.5	16.4	779	1 NIM1_NEUCR	P48479 neurospora
24	409.5	16.4	431	1 NRKA_TRYBB	Q08942 trypanosoma
25	403.5	16.1	431	1 NRKB_TRYBB	Q03428 trypanosoma
26	388	15.5	446	1 KCC1_YEAST	P27466 saccharomyc
27	382.5	15.3	1062	1 CC7_SCHPO	P41892 schizosacch
28	380.5	15.2	682	1 SNK_RAT	Q9r012 rattus norv
29	376.5	15.1	443	1 ST4_HUMAN	Q9v6e0 homo sapien
30	372.5	14.9	682	1 SNK_MOUSE	P53351 mus musculus
31	370.5	14.8	685	1 SNK_HUMAN	Q9nyy3 homo sapien
32	368.5	14.7	447	1 KCC2_YEAST	P29517 saccharomyc
33	360.5	14.4	543	1 CHK2_HUMAN	O96017 homo sapien

34 355.5 14.2 1576 1 YLK3\_CAEEL  
35 354 14.2 603 1 PLK1\_HUMAN  
36 352.5 14.1 1360 1 TNIK\_HUMAN  
37 352 14.1 968 1 STKA\_HUMAN  
38 349.5 14.0 1239 1 M4K1\_HUMAN  
39 349 14.0 966 1 STKA\_MOUSE  
40 346.5 13.9 805 1 FUSE\_DROME  
41 346 13.8 805 1 CHK2\_MOUSE  
42 345 13.8 1233 1 M4K1\_MOUSE  
43 344 13.8 603 1 PLK1\_RAT  
44 342 13.7 426 1 ST25\_HUMAN  
45 342 13.7 705 1 CDC5\_YEAST

ALIGNMENTS

RESULT 1  
NEK4\_HUMAN  
ID NEK4\_HUMAN STANDARD; PRT; 841 AA.  
AC P51957;  
DT 01-OCT-1996 (Rel. 34, Created)  
DT 01-OCT-1996 (Rel. 34, Last sequence update)  
DT 10-OCT-2003 (Rel. 42, Last annotation update)  
DE Serine/threonine-protein kinase Nek4 (EC 2.7.1.37) (NimaA-related  
protein kinase 4) (Serine/threonine protein kinase 2)  
DE (Serine/threonine-protein kinase NRK2).  
GN NEK4 OR STK2.  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
OX NCBI\_TaxID=9606;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC TISSUE=Breast;  
EX MEDLINE=94268838; PubMed=8208544;  
RA Levedakou E.N., He M., Baptist E.W., Craven R.J., Cance W.G.,  
RA Welcsh P.D., Simmons A., Naylor S.L., Leach R.J., Lewis T.B.,  
RA Bowcock A., Liu E.T.;  
RT "Two novel human serine/threonine kinases with homologies to the cell  
cycle regulating Xenopus MO15, and NIMA kinases: cloning and  
characterization of their expression pattern.";  
RL Oncogene 9:1977-1988(1994).  
CC -!- FUNCTION: Seems to act exclusively upon threonine residues (By  
similarity).  
CC -!- CATALYTIC ACTIVITY: ATP + a protein = ADP + a phosphoprotein.  
CC -!- SUBCELLULAR LOCATION: Nuclear (Probable).  
CC -!- TISSUE SPECIFICITY: HIGHEST EXPRESSION IN ADULT HEART, FOLLOWED BY  
PANCREAS, SKELETAL MUSCLE, BRAIN, LIVER, KIDNEY, LUNG AND  
PLACENTA, PRESENT IN MOST PRIMARY CARCINOMAS.  
CC -!- SIMILARITY: Belongs to the Ser/Thr family of protein kinases. NIMA  
subfamily.

-----  
This SWISS-PROT entry is copyright. It is produced through a collaboration  
between the Swiss Institute of Bioinformatics and the EMBL Outstation -  
the European Bioinformatics Institute. There are no restrictions on its  
use by non-profit institutions as long as its content is in no way  
modified and this statement is not removed. Usage by and for commercial  
entities requires a license agreement (see <http://www.isb-sib.ch/announce/>  
or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
-----

EMBL; L20321; AAA36658.1; --  
PIR; I78885; I78885.  
HSSP; Q63450; 1806.  
DR Genew; HGNC:11399; NEK4.  
DR GK; P51957; --  
DR MIM; 601959; --  
DR GO; GO:0004674; P:protein serine/threonine kinase activity; TAS.  
DR GO; GO:0006468; P:protein amino acid phosphorylation; TAS.  
DR InterPro; IPR000719; Prot\_kinase.  
DR InterPro; IPR008271; Ser\_Thr\_pkin\_AS.  
DR InterPro; IPR002290; Ser\_Thr\_pkinase.  
DR InterPro; IPR001245; Tyr\_pkinase.

DR Pfam; PF00069; pkinase; 1.  
 DR PRINTS; PR00109; TYRKINASE.  
 DR ProDom; PD000001; Prot kinase; 1.  
 DR SMART; SM00220; S TKC; 1.  
 DR PROSITE; PS00107; PROTEIN KINASE ATP; 1.  
 DR PROSITE; PS00108; PROTEIN KINASE ST; 1.  
 DR PROSITE; PS00111; PROTEIN KINASE ST; 1.  
 KW Transferase; Serine/threonine-protein kinase; ATP-binding; Mitosis;  
 KW Nuclear protein; Phosphorylation.  
 FT DOMAIN 6 261 PROTEIN KINASE.  
 FT NP BIND 12 20 ATP (BY SIMILARITY).  
 FT BINDING 35 35 ATP (BY SIMILARITY).  
 FT ACT\_SITE 131 131 BY SIMILARITY.  
 FT MOD\_RES 165 165 PHOSPHORYLATION (AUTO-) (BY SIMILARITY).  
 SQ SEQUENCE 841 AA; 94571 MW; 0DD31920DDE7EA58 CRC64;  
 Query Match 24.6%; Score 615; DB 1; Length 841;  
 Best Local Similarity 34.7%; Pred. No. 1e-24;  
 Matches 157; Conservative 87; Mismatches 152; Indels 56; Gaps 17;  
 QY 29 YVLQKLGSGSGFTGVYLVSDKXKAKGEEELKVLKBIISVGLNPNETVQANLEAQLLSKLDH 88  
 DB 6 YCYLRVVGSGSYGEVTLV---KHRRDGQYVTKLNLNASSRRERRAAEQEAQLLSQLKH 62  
 QY 89 PAIVKPHASFEVDNFC-IITEYCEGRDLDKIQEYKQAGKIFPENQIIEWFIQLLQVD 147  
 DB 63 PNIVTYKESWEGDGLLYVMGFCEGGDLYRKLKE--QKQQLPENQVWFVQIAMAQ 120  
 QY 148 YMHERRILRLDKSNVFL--KNLLKIGDFGVSRLLMGSCDLATLTGTTPHYMSPEALKH 206  
 DB 121 YLHKEHILRLDKTNQVFLTRNLIKVGLDGLIARVLEHNCDAVSTLIGTPPYMSPELFSN 180  
 QY 207 QGYDTSKDSIWSLACILYECMMHAFAGSNFLSIVLKVSGTPTSLPERYPKELNAIVES 266  
 DB 181 KPNYKSDYVWALGCCVYEWATLKHAFNAKMDMSLVYRIEGLKLPAMPDYSPELAELRT 240  
 QY 267 MLNKNPSLRPSAIEILKIPYLDQQLNLMCRYSEMLED--KNLDCQ-KEAAHIIN--- 319  
 DB 241 MLSKREPERPSVRSILRQYIKRQISFFLEATKIKTSKNKNGDSQSKPATVTVSGAE 300  
 QY 320 ANQRIHLQTLRA-----LSEVQKMPTRNRLRLK--QAADKARKLKIVVEKYEN 371  
 DB 301 SNHEVTHPQLSEGSQTYINGEKLCSQEKPRASGLLSPASLKAHCKQDL-----SN 355  
 QY 372 SKRMQBLRSRNFQQLSDVVLHKEHLKGMEE-----XEEQPE-----GRLSQSPQD 417  
 DB 356 TTELATISSV-----IDL-----PAKGRDSVSGFVQENQRYLDASNELGGI-CSISQ 405  
 QY 418 EDERWQGRBESDEFTLENLPESQIPSMOL 449  
 DB 406 VEEEMLDQNTKSAQP--ENL---IPMWSSDI 432

## RESULT 2

NER4\_MOUSE STANDARD; PRT; 792 AA.  
 AC Q9ZIJ2; O35673; Q9RIJ1;  
 DT 10-OCT-2003 (Rel. 42, Created)  
 DT 10-OCT-2003 (Rel. 42, Last sequence update)  
 DT 10-OCT-2003 (Rel. 42, Last annotation update)  
 DE Serine/threonine-protein kinase Nek4 (EC 2.7.1.37) (NIMA-related  
 DE protein kinase 4) (Serine/threonine protein kinase 2).  
 GN NEK4 OR STK2.  
 OS Mus musculus (Mouse).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 OX NCBI\_TaxID=10090;  
 RN [1]  
 RP SEQUENCE FROM N.A. (ISOFORMS 1 AND 2).  
 RC STRAIN=BALE/C; TISSUE=Spleen;  
 RX MEDLINE=20001940; PubMed=10529384;  
 RA Hayashi K., Igarashi H., Ogawa M., Sakaguchi N.;  
 RT "Activity and substrate specificity of the murine STK2

RT serine/threonine kinase that is structurally related to the mitotic  
 RT regulator protein NIMA of *Aspergillus nidulans*.";  
 RL Biochem. Biophys. Res. Commun. 264:449-456(1999).  
 RN [2]  
 RP SEQUENCE FROM N.A. (ISOFORM 1).  
 RX MEDLINE=9321807; PubMed=10393247;  
 RA Chen A., Yanai A., Arana E., Kilfin G., Motro B.;  
 RT "NIMA-related kinases: isolation and characterization of murine nek3  
 RT and nek4 cDNAs, and chromosomal localization of nek1, nek2 and nek3.";  
 RL Gene 234:127-137(1999).  
 CC -!- FUNCTION: Seems to act exclusively upon threonine residues.  
 CC -!- CATALYTIC ACTIVITY: ATP + a protein = ADP + a phosphoprotein.  
 CC -!- SUBCELLULAR LOCATION: Nuclear (Probable).  
 CC -!- ALTERNATIVE PRODUCTS:  
 CC Event=Alternative splicing; Named isoforms=2;  
 CC Name=1; Synonyms=mSTK2L;  
 CC IsoId=Q9ZIJ2-1; Sequence=Displayed;  
 CC Name=2; Synonyms=mSTK2S;  
 CC IsoId=Q9ZIJ2-2; Sequence=VSP\_007001;  
 CC -!- TISSUE SPECIFICITY: Expressed ubiquitously among various organs  
 CC and is upregulated in the testis.  
 CC -!- SIMILARITY: Belongs to the Ser/Thr family of protein kinases. NIMA  
 CC subfamily.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@sib-sib.ch](mailto:license@sib-sib.ch)).  
 CC -----  
 DR EMBL; AJ223071; CAAL1072.1; -.  
 DR EMBL; Y09234; CAA70436.1; -.  
 DR EMBL; AF090667; AAD16287.1; -.  
 DR PIR; JC7122; JC7122.  
 DR HSP; Q63450; IA06.  
 DR MGD; MGI:1344404; Nek4.  
 DR InterPro; IPR000719; Prot\_kinase.  
 DR InterPro; IPR08271; Ser\_Thr\_kinase.  
 DR InterPro; IPR002290; Ser\_Thr\_kinase.  
 DR InterPro; IPR001245; Tyr\_kinase.  
 DR Pfam; PF00069; pkinase; 1.  
 DR PRINTS; PR00109; TYRKINASE.  
 DR ProDom; PD000001; Prot kinase; 1.  
 DR SMART; SM00220; S TKC; 1.  
 DR PROSITE; PS00107; PROTEIN KINASE ATP; 1.  
 DR PROSITE; PS00111; PROTEIN KINASE ST; 1.  
 DR PROSITE; PS00108; PROTEIN KINASE ST; 1.  
 KW Transferase; Serine/threonine-protein kinase; ATP-binding; Mitosis;  
 KW Nuclear protein; Phosphorylation; Alternative splicing.  
 FT DOMAIN 6 261 PROTEIN KINASE.  
 FT NP BIND 12 20 ATP (BY SIMILARITY).  
 FT BINDING 35 35 ATP (BY SIMILARITY).  
 FT ACT\_SITE 131 131 BY SIMILARITY.  
 FT MOD\_RES 165 165 Missing (in isoform 2).  
 FT VARSPLIC 456 503 /FTID=VSP\_007001.  
 FT FT CONFLICT 499 499 G -> R (IN REF. 2).  
 SQ SEQUENCE 792 AA; 88994 MW; CF9187311C807A1D CRC64;  
 Query Match 24.0%; Score 600.5; DB 1; Length 792;  
 Best Local Similarity 31.5%; Pred. No. 5.4e-24;  
 Matches 141; Conservative 90; Mismatches 143; Indels 73; Gaps 13;  
 QY 29 YVLQKLGSGSGFTGVYLVSDKXKAKGEEELKVLKBIISVGLNPNETVQANLEAQLLSKLDH 88  
 DB 6 YCYLRVVGSGSYGEVTLV---KHRRDGQYVTKLNLNASSRRERRAAEQEAQLLSQLKH 62  
 QY 89 PAIVKPHASFEVDNFC-IITEYCEGRDLDKIQEYKQAGKIFPENQIIEWFIQLLQVD 147  
 DB 63 PNIVTYKESWEGDGLLYVMGFCEGGDLYRKLKE--QKQQLPENQVWFVQIAMAQ 120

QY 148 YMHERRILHRLDKSKNVFL-KNNLLKIGDFGVSRLLMGSCDLATTLTGTPHYMSPEALKH 206  
 Db 121 YLHEKHILHRLDKTQNVFLTRNIIKVGDLGIARVLEHNGDVASTLIGTPHYMSPELFSN 180  
 QY 207 QGYDTKSDIWSLACILYEMCCMHAFAGSNFLSIVLKIVGEGTSPSLPRYPKELNAIMES 266  
 Db 181 KPNYKSDVWALGCCVYEWATLKGFANAKDMNSLVYRIIEGKLPPMPKPVYSTELAEIART 240  
 QY 267 MLNKNPSLRPSAIEILKIPYLDEQL-----QNLNCRYSMTL-----ED 305  
 Db 241 MLSRRPEERPSVRSILRQYIKHHISLFLEATKAKTSKNNVNCDSRAKPAVAVSRKEE 300  
 QY 306 KULD-----CKEAAHIN-----AMQKRIHQLTALSEVQKMTPREMRURKQLQA 352  
 Db 301 SNTDVIHYQPRSESGSALVMGDKLSQEKFPVDIGPLRSPASLESGHTGKQDM-----NN 355  
 QY 353 ADEKARKLAKI-----VVEKYEENSKRMQELSRNFQOLSVDVLHEKTHLKGMEKEEQ 406  
 Db 356 TGESCATISIRINDILPAERDSANAGVQE-----SQPHVDADEVDSQ 401  
 QY 407 PEGRLJSCPDDEERQWQREESDEP 433  
 Db 402 -----CSISQE-KERLQGNTRSSDQ 421

RESULT 3

NEKI\_MOUSE  
 ID NEKI\_MOUSE STANDARD; PRT; 774 AA.  
 AC P51954;  
 DT 01-OCT-1996 (Rel. 34, Created)  
 DT 01-OCT-1996 (Rel. 34, Last sequence update)  
 DT 10-OCT-2003 (Rel. 42, Last annotation update)  
 DE Serine/threonine-protein kinase Nek1 (EC 2.7.1.37) (NimaA-related  
 DE protein kinase 1).  
 GN NEK1.  
 OS Mus musculus (Mouse).  
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 OX NCBI\_TaxID=10090;  
 RN [1]  
 RC SEQUENCE FROM N.A.  
 RC TISSUE=Blood;  
 RA MEDLINE=93010942; PubMed=1382974;  
 RA Letwin K., Mizzen L., Metro B., Ben-David Y., Bernstein A.,  
 RA Pawson T.;  
 RT "A mammalian dual specificity protein kinase, Nek1, is related to the  
 RT Nima cell cycle regulator and highly expressed in meiotic germ  
 RT cells.";  
 RL EMBO J. 11:3521-3531(1992).  
 CC -!- FUNCTION: PHOSPHORYLATES SERINES AND THREONINES, BUT ALSO APPEARS  
 CC -!- TO POSSESS TYROSINE KINASE ACTIVITY. IMPLICATED IN THE CONTROL OF  
 CC MEIOSIS.  
 CC -!- CATALYTIC ACTIVITY: ATP + a protein = ADP + a phosphoprotein.  
 CC -!- SUBCELLULAR LOCATION: Nuclear (Probable).  
 CC -!- TISSUE SPECIFICITY: PREDOMINANTLY IN TESTES (GERM CELLS AND  
 CC SERTOLI CELLS). LOWER LEVELS IN OVARY (OVOCYTES AND GRANULOSA  
 CC CELLS), THYMUS, AND LUNG.  
 CC -!- DEVELOPMENTAL STAGE: IN FEMALE, EXPRESSED AS FOLLICLES ENTER THE  
 CC SECONDARY STAGE UNTIL OVULATION OCCURS. IN THE MALE REPRODUCTIVE  
 CC SYSTEM, THE EXPRESSION IS LIMITED TO SPERMATOCYTES AND SPERMATIDS.  
 CC -!- SIMILARITY: Belongs to the Ser/Thr family of protein kinases. NIMA  
 CC subfamily.

-----  
 This SWISS-PROT entry is copyright. It is produced through a collaboration  
 between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 the European Bioinformatics Institute. There are no restrictions on its  
 use by non-profit institutions as long as its content is in no way  
 modified and this statement is not removed. Usage by and for commercial  
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 -----  
 DE EMBL; S45828; AAB23529.1; -.  
 DE PIR; S25284; S25284.

HSSP; Q63450; 1A06.  
 MGD; MGI:97303; Nek1.  
 InterPro; IPR000719; Prot\_kinase.  
 InterPro; IPR008271; Ser\_thr\_kin\_As.  
 InterPro; IPR002290; Ser\_thr\_kinase.  
 InterPro; IPR001245; Tyr\_kinase.  
 Pfam; PF00069; pkinase; 1.  
 PRINTS; PR00109; TYRKINASE.  
 ProDom; PD000001; Prot\_kinase; 1.  
 SMART; SM00220; S\_TKC\_1.  
 PROSITE; PS0107; PROTEIN\_KINASE\_ATP; 1.  
 PROSITE; PS0108; PROTEIN\_KINASE\_ST; 1.  
 PROSITE; PS0011; PROTEIN\_KINASE\_DOM; 1.  
 KX Transferase; Serine/threonine-protein kinase; ATP-binding; Mitosis;  
 KW Nuclear protein; Phosphorylation; Cell cycle; Cell division;  
 KW Tyrosine-protein kinase.  
 FT DOMAIN 4 258 PROTEIN\_KINASE.  
 FT NP\_BIND 10 18 ATP (BY SIMILARITY).  
 FT BINDING 33 33 ATP (BY SIMILARITY).  
 FT ACT\_SITE 128 128 BY SIMILARITY.  
 FT MOD\_RES 162 162 PHOSPHORYLATION (AUTO-) (BY SIMILARITY).  
 SQ SEQUENCE 774 AA; 88427 MW; 7EDES881ACDC06FA CRC64;  
 Query Match 23.9%; Score 598; DB 1; Length 774;  
 Best Local Similarity 28.5%; Pred.No. 7.1e-24;  
 Matches 141; Conservative 94; Mismatches 132; Indels 128; Gaps 13;  
 QY 28 RYVLQKLGSGSFGTVYVLSVSKAKGGEELKVLKEISVGEINPNETVQANLEAQLLSKLD 87  
 Db 3 KYVRLQKIGSGFGKAVLV--KSTEDGRHY-VIKEINISRMDSKQERESRREAVLANMK 59  
 QY 88 HPAIVKHFASFVEQDNFCIITEYCEGDDDDKQYKQAGKIPENQIIEFFQLLLGVD 147  
 Db 60 HPNTVQKESFENGSLYIVNDYCEGGDLFKRIN--AQKALFQEDQILDWFVQICLALK 117  
 QY 148 YMHERRILHRLDKSKNVFL-KNNLLKIGDFGVSRLLMGSCDLATTLTGTPHYMSPEALKH 206  
 Db 118 YVHDKILHRLDIKSONIFLKDCGTQVQLGDFGIARVLEHNGDVASTLIGTPHYLSPEICEN 177  
 QY 207 QGYDTKSDIWSLACILYEMCCMHAFAGSNFLSIVLKIVGEGTSPSLPRYPKELNAIMES 266  
 Db 178 KPNYKSDIWSLACILYELCTLHAFEGAGNNKVLKIIISGSPFPVSPHYSDLRSLSSQ 237  
 QY 267 MLNKNPSLRPSAIEILKIPYLDEQLNLM-----C----- 296  
 Db 238 LFRKNPRDRFSVNSILEKGFIAKRIEFLSPQIAEFCULTKSKFGPQLPGKRPASGQ 297  
 QY 297 -----RYSEMTLDRKNDLCOKEAAHIINAMQKRIHLQTLR 331  
 Db 298 GVSSFVPAQKITPAKAYGVPLATYKYDKKLEKKPPPKHQAQI----- 344  
 QY 332 ALSVQVQKTPREMRURLEKLOADEKARKLKIIVEEKYEENSKRMQELSRNFQOLSVDVL 391  
 Db 345 ---PVKQNSGEERKKMSEEA--KKRLEFIEKEK-----KOKDQIRFLKAEQMK---R 391  
 QY 392 HEKTHLKGMEKEEQP-----EGRILS---CSPQ----- 416  
 Db 392 QEQRLERINRAREQGRNVLRAGSGEVKASFFGIGGAVSPSPSPGQYEHVHAFDQ 451  
 QY 417 -----DEDERWQ 425  
 Db 452 MQLRAEDNEARWKG 466  
 RESULT 4  
 NEK3\_HUMAN  
 ID NEK3\_HUMAN STANDARD; PRT; 506 AA.  
 AC P51956; Q8WUN5;  
 DT 01-OCT-1996 (Rel. 34, Created)  
 DT 28-FEB-2003 (Rel. 41, Last sequence update)  
 DT 15-MAR-2004 (Rel. 43, Last annotation update)  
 DE Serine/threonine-protein kinase Nek3 (EC 2.7.1.37) (NimaA-related  
 DE protein kinase 3) (HSPK 36).





use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announcement/> or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

EMBL; AB067488; BAB67794.1; ALT\_INIT.  
EMBL; AF155113; AAD42879.1; --  
Genew; HGNC:7744; NEK1.  
GK; Q96PY6; --  
MIM; 604588; --  
InterPro; IPR000719; prot kinase.  
InterPro; IPR008271; Ser Thr pkin\_AS.  
InterPro; IPR002290; Ser Thr pkinase.  
InterPro; IPR001245; Tyr pkinase.  
Pfam; PF00069; pkinase; 1.  
PRINTS; PR00109; TYRKINASE.  
ProDom; PD000001; Prot kinase; 1.  
SMART; SM00220; S TKC; 1.  
PROSITE; PS00107; PROTEIN KINASE ATP; 1.  
PROSITE; PS0011; PROTEIN KINASE\_DOM; 1.  
PROSITE; PS00108; PROTEIN KINASE\_ST; 1.  
Transferase; Serine/threonine-protein kinase; ATP-binding; Mitosis;  
Nuclear protein; Phosphorylation; Cell cycle; Cell division;  
Tyrosine-protein kinase; Alternative splicing.  
DOMAIN 4 258 PROTEIN KINASE.  
NP\_BIND 10 18 ATP (BY SIMILARITY).  
BINDING 33 33 ATP (BY SIMILARITY).  
ACT\_SITE 128 128 BY SIMILARITY.  
MOD\_RES 162 162 PHOSPHORYLATION (AUTO-) (BY SIMILARITY).  
VARSPPLIC 478 521 MISSING (in isoform 2).  
FTID=VSP 004870.  
CONFLICT 1232 1232 G -> E (IN REF. 2).  
SEQUENCE 1258 AA; 142828 MW; 339C4BFA56612530 CRC64;

Query Match 23.3%; Score 582.5; DB 1; Length 1258;  
Best Local Similarity 29.9%; Pred. No. 7.6e-23;  
Matches 122; Conservative 93; Mismatches 124; Indels 69; Gaps 8;  
28 RYVLOKLGSGSGTGYLYSDKAKRGEELKVLKETSVCGLNPNFVQANLQALLSKLD 87  
3 KYRVLQKIGEGSGKAILV--KSTEDGRQY-VIKELINISMSKREESRREAVLANWK 59  
88 HPAIVKFAFVQDNFCITTEYCEGRDLDKIQEYKQAGKIFPNQIIEWTQILLGV 147  
60 HPNIVQYRESFEENGSLYIVMDYCEGDLFKRIN--AQKGVLFQEDQILDWVQICLALK 117  
148 YMERILHDLKSKVEL-KNNLLKIGDGVSRLLMGSCDLATTLTGTPHYMSPEALKH 206  
118 HVHDKILHARDIKSNIFKTOGTQVDFGFLARVLNSTVELARTICIGIPYLSPEICEN 177  
207 QGYDTKSDIWSLACILYEMCMNFAFAGNFIISVLKIVEGDTSPSPERYKELNAIMES 266  
178 KPTNKSDDIHALGCVLYELCTLKHAEAGSMKVLKLIISGSPFPPVSLHYSDLRSLVSQ 237  
267 MLNKNLSRPSAETILKIPYLDQQLNLM-----C----- 296  
238 LFKRNPDRPSVNSILEKGFIAKRIKFLSPQIAEEFLKTFKFSQSPIPAKRPASQ 297  
297 -----RYSEMTEDKNDLQCKEAAHINAMQRIHLQTLR 331  
298 NSISVMPAQKITPAKYGIPLAYKYKYGKGLHEKKLPQKHQAH--QPFKKVN----- 350  
332 ALSEVQMTFRNRLKLOAADEKARKLKKIVEEKYENSKMQELR 379  
351 -TGERRKISEEARKLERLFIEKKKQKQDQIISLMKASQMKRQEKER 397

Query Match 23.3%; Score 582.5; DB 1; Length 1258;  
Best Local Similarity 29.9%; Pred. No. 7.6e-23;  
Matches 122; Conservative 93; Mismatches 124; Indels 69; Gaps 8;

Query Match 23.3%; Score 582.5; DB 1; Length 1258;  
Best Local Similarity 29.9%; Pred. No. 7.6e-23;  
Matches 122; Conservative 93; Mismatches 124; Indels 69; Gaps 8;

15-MAR-2004 (Rel. 43, Last annotation update)  
Serine/threonine-protein kinase Nek2 (EC 2.7.1.37) (NIMA-related protein kinase 2) (NIMA-like protein kinase 1) (HSPK 21).  
NEK2 OR NLK1.  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
OX NCBI\_TaxID=9606;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC TISSUE=Nasopharynx, Placenta, and T-cell;  
RX MEDLINE=94368699; PubMed=7522034;  
RA Schultz S.J., Fry A.M., Suetterlin C., Ried T., Nigg E.A.;  
RT "Cell cycle-dependent expression of Nek2, a novel human protein kinase related to the NIMA mitotic regulator of Aspergillus nidulans";  
RL Cell Growth Differ. 5:625-635(1994).  
RN [2]  
RP SEQUENCE FROM N.A.  
RC Lu K.P., Hunter T.;  
RT "Molecular cloning and expression of NLK1, a human NIMA-like kinase";  
RL Submitted (JUL-1994) to the EMBL/GenBank/DBJ databases.  
RN [3]  
RP SEQUENCE FROM N.A.  
RC TISSUE=Skin;  
RX MEDLINE=22388257; PubMed=12477932;  
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,  
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,  
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,  
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,  
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,  
RA Stapleton M.J., Udell T.B., Toshiyuki S., Carninci P., Prange C.,  
RA Brownstein M.J., Peters G.J., Abramson R.D., Mullaly S.J.,  
RA Raba S.S., Loquellano N.A., Casavant T.L., Scheetz T.E.,  
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,  
RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,  
RA Fahey J., Helton E., Kettelman M., Maman A., Rodriguez S., Sanchez A.,  
RA Whiting M., Maman A., Young A.C., Shevchenko Y., Bouffard G.G.,  
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,  
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,  
RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smallos D.E.,  
RA Scherch A., Schein J.E., Jones S.J.M., Marra M.A.;  
RT "Generation and initial analysis of more than 15,000 full-length human and mouse cDNA sequences";  
RL Proc. Natl. Acad. Sci. U.S.A. 99:16999-16903(2002).  
RN [4]  
RP SEQUENCE OF 83-203 FROM N.A.  
RX MEDLINE=94100173; PubMed=8274451;  
RA Schultz S.J., Nigg E.A.;  
RT "Identification of 21 novel human protein kinases, including 3 members of a family related to the cell cycle regulator nima of Aspergillus nidulans";  
RL Cell Growth Differ. 4:821-830(1993).  
CC -!- FUNCTION: PROTEIN KINASE THAT IS INVOLVED IN MITOTIC REGULATION.  
CC MAY HAVE A ROLE AT THE G2-M TRANSITION. MAY ALSO PLAY A ROLE IN MEIOSIS.  
CC -!- CATALYTIC ACTIVITY: ATP + a protein = ADP + a phosphoprotein.  
CC -!- SUBUNIT: Interacts with TERF1.  
CC -!- SUBCELLULAR LOCATION: Nuclear (By similarity).  
CC -!- DEVELOPMENTAL STAGE: ACCUMULATES THROUGHOUT S PHASE AND SHOWS MAXIMAL LEVELS IN LATE G2. THIS EXPRESSION PATTERN IS HIGHLY REMINISCENT OF THAT OF A AND B CYCLINS.  
CC -!- SIMILARITY: Belongs to the Ser/Thr family of protein kinases. NIMA subfamily.  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announcement/> or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).



```
CC DR EMBL; Z29066; CAA82309.1; -
DR EMBL; U11050; AAA19558.1; -
DR EMBL; BC043502; AAH43502.2; -
DR EMBL; Z25425; CAA80912.1; -
DR PIR; G01452; G01452.
DR PIR; I38215; I38215.
DR HSSP; Q00534; IBI7.
DR Genew; HGNC:7745; NEK2.
DR GK; P51955; -.
DR MIM; 604043; -.
DR GO; GO:0005813; C:centrosome; TAS.
DR GO; GO:0004674; F:protein serine/threonine kinase activity; TAS.
DR GO; GO:0007088; P:regulation of mitosis; TAS.
DR InterPro; IPR000719; Prot_kinase.
DR InterPro; IPR008271; Ser_thr_pkin_AS.
DR InterPro; IPR002290; Ser_thr_pkinase.
DR Pfam; PF00069; pkinase; 1.
DR ProDom; PD000001; Prot_kinase; 1.
DR SMART; SM00220; S_TK; 1.
DR PROSITE; PS00107; PROTEIN_KINASE_ATP; FALSE_NEG.
DR PROSITE; PS00108; PROTEIN_KINASE_ST; 1.
DR PROSITE; PS00011; PROTEIN_KINASE_DOM; 1.
DR PROSITE; PS00012; PROTEIN_KINASE_ST; 1.
DR TRANSFERASE; Serine/threonine-protein kinase; ATP-binding; Mitosis;
KW Nuclear protein; Phosphorylation; Cell cycle; Cell division; Meiosis.
FT DOMAIN 8 271 PROTEIN_KINASE.
FT NP_BIND 14 22 ATP (BY SIMILARITY).
FT BINDING 37 37 ATP (BY SIMILARITY).
FT ACT_SITE 141 141 BY SIMILARITY.
FT MOD_RES 175 175 PHOSPHORYLATION (AUTO-) (BY SIMILARITY).
FT CONFLICT 84 85 IV -> LY (IN REF. 4).
SQ SEQUENCE 445 AA; 51763 MW; D33A37778ABB6D9E CRC64;

Query Match
Best Local Similarity 19.8%; Score 495; DB 1; Length 445;
Matches 139; Conservative 70; Mismatches 168; Indels 48; Gaps 14;

QY 26 ARYVLQKLGSGSFGTVYLVSDKKAKGEEELKVL--KEISVGEINPNETVQANLEAQL 83
Db 5 AEDYEVLVITGTSGYRC-----QKIRKSDGKILVWKELDYGSTAEKQMLVSEVNL 59
QY 84 SKLDHPAIVKFKASVEQDN--FCILTEYCEGRDLDDKIOEQKAGKIPPEINQIETW 141
Db 60 RELKHPNIVRYDRIIDRTNTLYIMVEYCEGGDLASVITKTKEROYLDEEFLVRVMTQ 119
QY 142 LLLGVVDYMER-----RIHLRLKSNVPL--KNLLKIGDFGSRLLMGSCDLATTLTG 194
Db 120 LTLALKECHRRSDGGHTVLRDLKANVFLDGKQN-VKLGDFGLARILNHDTSFAKITVG 178
QY 195 TPHYSPALKHQGYDTKSDIWSLACILYEMCOMHAFAGSNFLSIVLKEGDTPLPE 254
Db 179 TPYMSPEQMNRMSYNEKSDIWSLGLLYELCALMPPTAFSQELAGKIREGKFRIPY 238
QY 255 RYPKEINAIMESMLKNPSLRPSAIEILKIPVL-----DEQLQNLWCYSEMTLEDK 309
Db 239 RYSDENELITRLNLKDYHRSVEIENPLDIADLVADEQRNLRGRQLGEPKESQD 298
QY 310 COKEAHIINAVQ-KRIHLQ-----TIRALSE-----VQMTPRRMRLKQLQADEK 359
Db 299 ----SSPVLSELKKEIQLOERERALKAREERLEQKEQLCVRLAEEDKLARENLLKN 354
QY 360 LKKIIVEEYKESKMQELRSNFQSLVDVLHEKTHLKGWEEK-----EQPEGRSLCS 414
Db 355 YSLKERKFLSLASNPPELL-----NLPSSVITKKVHFSG-ESKENIMRSENSSQLTSK 407
QY 415 PODED 419
Db 408 SKCKD 412

RESULT 8
NEK8_BRARE
ID NEK8_BRARE STANDARD; PRT; 697 AA.
```

```
AC Q90XC2;
DT 10-OCT-2003 (Rel. 42, Created)
DT 10-OCT-2003 (Rel. 42, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Serine/threonine-protein kinase NEK8 (EC 2.7.1.37) (NimaA-related
DE protein kinase 8).
GN NEK8.
OS Brachydanio rerio (Zebrafish) (Danio rerio).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes;
OC Cyprinidae; Danio.
OX NCBI_TaxID=7955;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=22308894; PubMed=12421721;
RA Liu S., Lu W., Obara T., Kuida S., Lehoczy J., Dewar K.,
RA Drummond I.A., Beier D.R.;
RT "A defect in a novel Nek-family kinase causes cystic kidney disease in
RT the mouse and in zebrafish.";
RL Development 129:5839-5846(2002).
CC -!- FUNCTION: Required for renal tubular integrity.
CC -!- CATALYTIC ACTIVITY: ATP + a protein = ADP + a phosphoprotein.
CC -!- SUBCELLULAR LOCATION: Predominantly cytoplasmic (By similarity).
CC -!- DISEASE: Defects in NEK8 are the cause of polycystic kidney
CC disease.
CC -!- SIMILARITY: Belongs to the Ser/Thr family of protein kinases. NIMA
CC subfamily.
CC -!- SIMILARITY: Contains 5 RCK1 repeats.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@sib-sib.ch).
CC
CC -----
CC EMBL; AF407580; AAL09676.1; -.
CC ZFIN; ZDB-GENE-020509.1; nek8.
CC InterPro; IPR000719; Prot_kinase.
CC InterPro; IPR000408; Reg_Chrr_condens.
CC InterPro; IPR008271; Ser_thr_pkin_AS.
CC InterPro; IPR002290; Ser_thr_pkinase.
CC InterPro; IPR001245; Tyr_pkinase.
CC Pfam; PF00069; pkinase; 1.
CC Pfam; PF00415; RCK1; 3.
CC PRINTS; PR00633; RCKNSATION.
CC PRINTS; PR00109; TYRKINASE.
CC ProDom; PD000001; Prot_kinase; 1.
CC SMART; SM00220; S_TK; 1.
CC PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.
CC PROSITE; PS00011; PROTEIN_KINASE_DOM; 1.
CC PROSITE; PS00108; PROTEIN_KINASE_ST; 1.
CC PROSITE; PS00625; RCK1_1; FALSE_NEG.
CC PROSITE; PS00626; RCK1_2; FALSE_NEG.
CC PROSITE; PS00012; RCK1_3; 4.
CC Transferase; Serine/threonine-protein kinase; ATP-binding; Mitosis;
KW Phosphorylation; Repeat.
KW DOMAIN 4 263 PROTEIN_KINASE.
FT NP_BIND 10 18 ATP (BY SIMILARITY).
FT REPEAT 417 468 RCK1 1.
FT REPEAT 469 520 RCK1 2.
FT REPEAT 521 586 RCK1 3.
FT REPEAT 587 636 RCK1 4.
FT REPEAT 637 689 RCK1 5.
FT ACT_SITE 128 128 BY SIMILARITY.
FT BINDING 33 33 ATP (BY SIMILARITY).
FT MOD_RES 162 162 PHOSPHORYLATION (AUTO-) (BY SIMILARITY).
SQ SEQUENCE 697 AA; 76541 MW; 875A51D1F3831AA7 CRC64;

Query Match
Best Local Similarity 19.0%; Score 474.5; DB 1; Length 697;
Matches 111; Conservative 63; Mismatches 107; Indels 31; Gaps 7;
```



FT BINDING 33 33 ATP (BY SIMILARITY).  
FT MOD\_RES 162 162 PHOSPHORYLATION (AUTO-) (BY SIMILARITY).  
SQ SEQUENCE 692 AA; 74806 MW; 9E09820DFB3D5CA1 CRC64;

Query Match 18.2%; Score 454; DB 1; Length 692;  
Best local similarity 36.1%; Pred. No. 1.5e-16;  
Matches 97; Conservative 61; Mismatches 103; Indels 8; Gaps 4

QY 28 RYVLOOKLGSGFGFVYLVSDDKAKRGEELVKLEISVGELPNPNETVQANLAQLLSKLD 87  
DB :|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:  
3 KYERIVRVGRGAGIVHLCL--LRKADQKLVIKIQIPVQMTEKEERQAQNCEQVLKLN 59  
QY 88 HPAIVKFHASVFQDNFCITTEYCEGRDLDDTKIOEQKQAGKIFPENQIIENFIQLLLGVD 147  
DB :|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:  
60 HENVIEYVENFEDEKALMIAMAPCGTIAEFIQ--KRCNSLIESTILHFVEVQILLAH 117  
QY 148 YMHERRIILHRDLKSNVFLKN--LLKIGDFGVSRLLMGSCDLATTLTGTPHYMSPEALK 205  
DB :|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:  
118 HVHTHLLIHRDLKTQNLIDKHNVKIKGFEGSKIL-SKSAYTVVGTGPYISPELCE 176  
QY 206 HGQYPTKSDIWSLACILIYEMCMNHPAGSNFLSIVLKIVEGDTPSLPERYPKELNAIME 265  
DB :|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:  
177 GKPYNQKSDIWALGCVLVELASLKRAFEAANPALVKIMSGTFAPISDRYSPELRQLVL 236  
QY 266 SWLNKNPSLRPSAIEILKIPVDELQONL 294  
DB :|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:  
237 LLSLEPAQRPLSHINVAQPLCTRALNLL 265

RESULT 11  
NEK8 MOUSE  
ID\_NEK8\_MOUSE STANDARD; PRT; 698 AA.  
AC Q91ZE4; Q9D695;  
DT 10-OCT-2003 (Rel. 42, Created)  
DT 10-OCT-2003 (Rel. 42, Last sequence update)  
DT 10-OCT-2003 (Rel. 42, Last annotation update)  
DE Serine/threonine-protein kinase Nex8 (EC 2.7.1.37) (NimA-related protein kinase 8).  
GN NEK8  
OS Mus musculus (Mouse).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
NCBI\_TaxId=10090;  
[1]  
RP SEQUENCE FROM N.A. (ISOFORM 1), MUTAGENESIS OF LYS-33, AND VARIANT ARJPKD VAL->448.  
RP STRAIN=C57BL/6J;  
RX MEDLINE=22308894; PubMed=12421721;  
RA Liu S., Lu W., Obara T., Kuida S., Lehoczy J., Dewar K., Drummond I.A., Beier D.R.;  
RA "A defect in a novel Nek-family kinase causes cystic kidney disease in the mouse and in zebrafish."  
RL Development 129:5839-5846(2002).  
RN [2]  
RP SEQUENCE FROM N.A. (ISOFORM 2).  
RP STRAIN=C57BL/6J; TISSUE=Skin;  
RX MEDLINE=21085660; PubMed=11217851;  
RA Kawai J., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y., Arakawa T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S., Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamataka I., Saio T., Okazaki Y., Gojobori T., Bono H., Kasukawa T., Saio R., Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T., Fleischmann W., Gaasterland T., Gissi C., King B., Kochiya H., Khel P., Lewis S., Matsuo Y., Nikola I., Pesole G., Quackenbush J., Schriml L.M., Szaubli P., Suzuki R., Tomita M., Wagner L., Washio T., Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G., Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F., Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M., Gustincich S., Hill D., Hofmann C., Hume D.A., Kamiya M., Lee N.H., Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaerts P., Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N., Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-P., Suzuki H., Toyo-oka K., Wang K.H., Weitz C., Whitaker C., Wilming L.

Wynshaw-Boris A., Yoshida K., Hasegawa Y., Kawaji H., Kohsaki S., Hayashizaki Y.;  
 "Functional annotation of a full-length mouse cDNA collection."  
 Nature 409:685-690(2001).  
 CC -!- FUNCTION: Required for renal tubular integrity. May regulate local cytoskeletal structure in kidney tubule epithelial cells.  
 CC -!- CATALYTIC ACTIVITY: ATP + a protein = ADP + a phosphoprotein.  
 CC -!- SUBCELLULAR LOCATION: Predominantly cytoplasmic.  
 CC -!- ALTERNATIVE PRODUCTS:  
 CC Event=Alternative splicing; Named isoforms=2;  
 CC Name=1;  
 CC IsoId=Q91ZR4-1; Sequence=Displayed;  
 CC Name=2;  
 CC IsoId=Q91ZR4-2; Sequence=VSP\_007959, VSP\_007960;  
 CC Note=Due to intron retention. No experimental confirmation available;  
 CC -!- TISSUE SPECIFICITY: Kidney, liver, and testis.  
 CC -!- DISEASE: Defects in NEK8 are the cause of autosomal recessive juvenile polycystic kidney disease (ARJPKD).  
 CC -!- SIMILARITY: Belongs to the Ser/Thr family of protein kinases. NIMA subfamily.  
 CC -!- SIMILARITY: Contains 5 RCC1 repeats.  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 CC EMBL: AF407579; AAL09675.1; -;  
 CC EMBL: AK014546; BAB29424.1; -;  
 CC MGI: MG1:1890646; Nek8.  
 CC InterPro: IPR000719; Prot\_kinase.  
 CC InterPro: IPR000408; Reg\_chr\_condens.  
 CC InterPro: IPR008271; Ser\_thr\_pkin\_AS.  
 CC InterPro: IPR002290; Ser\_thr\_pkinase.  
 CC InterPro: IPR001245; Tyr\_pkinase.  
 CC Pfam: PF00069; pkinase; 1.  
 CC Pfam: PF00415; RCC1; 3.  
 CC PRINTS: PR00633; RCCNDNSATION.  
 CC PRINTS: PR00109; TYRKINASE.  
 CC ProDom: PD000001; Prot\_kinase; 1.  
 CC SMART: SM00220; S\_TKc; 1.  
 CC PROSITE: PS00107; PROTEIN\_KINASE\_ATP; 1.  
 CC PROSITE: PS00011; PROTEIN\_KINASE\_DOM; 1.  
 CC PROSITE: PS00108; PROTEIN\_KINASE\_ST; 1.  
 CC PROSITE: PS00625; RCC1\_1; FALSE\_NEG.  
 CC PROSITE: PS00625; RCC1\_1; FALSE\_NEG.  
 CC PROSITE: PS00626; RCC1\_2; FALSE\_NEG.  
 CC PROSITE: PS0012; RCC1\_3; 5.  
 CC Transferrase; Serine/threonine-protein kinase; ATP-binding; Phosphorylation; Repeat; Alternative splicing; Disease mutation.  
 CC DOMAIN 4 258 PROTEIN\_KINASE.  
 CC NE\_BIND 10 18 ATP (BY SIMILARITY).  
 CC REPEAT 416 467 RCC1 1.  
 CC REPEAT 468 519 RCC1 2.  
 CC REPEAT 520 585 RCC1 3.  
 CC REPEAT 586 637 RCC1 4.  
 CC REPEAT 638 690 RCC1 5.  
 CC BY SIMILARITY.  
 CC ACT\_SITE 128 128 ATP (BY SIMILARITY).  
 CC BINDING 33 33 PHOSPHORYLATION (AUTO-) (BY SIMILARITY).  
 CC MOD\_RES 162 162 AEKSLTFGPPIAGS -> PVQDGSNGGHPVPT (in isoform 2).  
 CC VARSPPLIC 277 291 /FTid=VSP\_007959.  
 CC Missing (in isoform 2).  
 CC FTid=VSP\_007960.  
 CC G->V (in ARJPKD).  
 CC K-M: RESULTS IN ENLARGED MULTINUCLEATED CELLS.  
 CC F->L (IN REF. 2).  
 CC R->G (IN REF. 2).  
 CC

SQ SEQUENCE 698 AA; 75264 MW; 686B29A8CF180E94 CRC64;  
 Query Match 18.2%; Score 454; DB 1; Length 698;  
 Best Local Similarity 36.1%; Pred. No. 1.6e-16;  
 Matches 97; Conservative 61; Mismatches 103; Indels 8; Gaps 4;  
 QY 28 RYVLOQKLGSGSGTGYLVSDKAKRGEELKVLKEISVGLNPNETVQANLEAQLLSKLD 97  
 DB 3 KYERIRVVGGAAGIVHLC---LRKADQKLVILKQIPVEQMTKEERQAQNECVLKLIN 59  
 QY 88 HPAIVKPHASFEVDNFCITTEYCEGRDLDKIQEYKQAGKIPFENQIETWITQLLGYD 147  
 DB 60 HPNVIEYENFLDKALMIAMEYAPGGTIAEFTQ--KRCNSLLEETILHFFVQILLALH 117  
 QY 148 YMHERRILHRDLKSNVFLKNN--LLKIGDFGVSRLLMGSCDLATLTGTGPHYMSPEALK 205  
 DB 118 HVHTHLILHRDLKTQNLIDKHMVVKIGDFGISKIL--SSKAYTVVGTGTCYISPCLCE 176  
 QY 206 HQGYTKSDIWSIACILYEMCMNHAPAGSNFISIVLKIIVEGTPSLPDPYKPELNAIME 265  
 DB 177 GKPNQKSDIWDALGCVLYELASLKRAFEAANLPALVKIMSGTFAPISDRYSPELRQLVL 236  
 QY 266 SMLNKPSPSLRPSAIEILKIPYLDQQLNL 294  
 DB 237 SLLSLEPAQRPPLSHIMAQPLCITRALINI 265  
 RESULT 12  
 NEK2\_MOUSE STANDARD; PRT; 443 AA.  
 ID NEK2\_MOUSE AC 035942; O35959;  
 DT 16-OCT-2001 (Rel. 40, Created)  
 DT 16-OCT-2001 (Rel. 40, Last sequence update)  
 DT 10-OCT-2003 (Rel. 42, Last annotation update)  
 DE Serine/threonine-protein kinase Nek2 (EC 2.7.1.37) (NIMA-related protein kinase 2).  
 GN NEK2.  
 OS Mus musculus (Mouse).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 OX NCBI\_TaxID=10090;  
 RN [1]  
 RP SEQUENCE FROM N.A., SUBCELLULAR LOCATION, AND TISSUE SPECIFICITY.  
 RC STRAIN=Swiss Webster; TISSUE=Testis;  
 RX MEDLINE=97330684; PubMed=9181143;  
 RA Rhee K., Wolgemuth D.J.;  
 RT "The NIMA-related kinase 2, Nek2, is expressed in specific stages of the meiotic cell cycle and associates with meiotic chromosomes.";  
 RN Development 124:2167-2177(1997).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=98243037; PubMed=9583679;  
 RA Arama E., Yanai A., Kilfin G., Motro B.;  
 RT "Murine NIMA-related kinases are expressed in patterns suggesting distinct functions in gametogenesis and a role in the nervous system.";  
 RN Oncogene 16:1813-1823(1998).  
 RN [3]  
 RP SEQUENCE FROM N.A., AND TISSUE SPECIFICITY.  
 RC TISSUE=Brain;  
 RX MEDLINE=98096235; PubMed=9434622;  
 RA Tanaka K., Parvainen M., Nigg E.A.;  
 RT "The in vivo expression pattern of mouse Nek2, a NIMA-related kinase, indicates a role in both mitosis and meiosis.";  
 RN Exp. Cell Res. 237:264-274(1997).  
 CC -!- FUNCTION: PROTEIN KINASE THAT IS INVOLVED IN MITOTIC REGULATION. MAY HAVE A ROLE AT THE G2-M TRANSITION. MAY ALSO PLAY A ROLE IN MEIOSIS.  
 CC -!- CATALYTIC ACTIVITY: ATP + a protein = ADP + a phosphoprotein.  
 CC -!- SUBUNIT: Interacts with TERF1 (by similarity).  
 CC -!- SUBCELLULAR LOCATION: Nuclear.  
 CC -!- TISSUE SPECIFICITY: MOST ABUNDANTLY EXPRESSED IN TESTIS. LOW LEVELS FOUND IN MID-GESTATION EMBRYO, OVARY, PLACENTA, INTESTINE.

```

RESULT 13
FIN1 SCHFO
ID ID SCHPO STANDARD; PRT; 722 AA.
AC O13939.
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE G2-specific protein kinase finl (EC 2.7.1.-).
DE FIN1 OR SPAC19E9.02.
OS Schizosaccharomyces pombe (Fission yeast).
OC Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;
OC Schizosaccharomycetales; Schizosaccharomycetaceae;
OC Schizosaccharomycetes.
OX NCBI_TaxID=4896;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=972;
RC MEDLINE=21848401; PubMed=11859360;
RA Wood V., Gilliam R., Rajandream M.A., Lyne M., Lyne R., Stewart A.,
RA Sgouros J., Peat N., Hayes J., Baker S., Basham D., Bowman S.,
RA Brooks K., Brown D., Brown T., Chillingworth T., Church C.M.,
RA Collins M., Connor R., Cronin A., Davis P., Fiddell T., Fraser A.,
RA Gentles S., Goble A., Hamlin N., Harris D., Hidalgo J., Hodgson G.,
RA Holroyd S., Hornsby T., Howarth S., Huckle E.J., Hunt S., Jagels K.,
RA James K., Jones L., Jones M., Leather S., McDonald S., McLean J.,
RA Mooney P., Moule S., Mungall K., Murphy L., Niblett D., Odell C.,
RA Oliver K., O'Neill S., Pearson D., Quail M.A., Rabinowitsch E.,
RA Rutherford K., Rutter S., Saunders D., Seeger K., Sharp S.,
RA Skelton J., Simmonds M., Squares R., Squares S., Stevens K.,
RA Taylor K., Taylor R.G., Tivey A., Walsh S.V., Warren T., Whitehead S.,
RA Woodward J., Volckaert G., Aert R., Robben J., Grymonprez B.,
RA Weljens I., Vanstraelen E., Rieger M., Schaefer M., Mueller-Auer S.,
RA Gabel C., Fuchs M., Fritzc C., Holzer E., Moestl D., Hilbert H.,
RA Borzym K., Langer I., Beck A., Lehrach H., Reinhardt R., Pohl T.M.,
RA Eger P., Zimmermann W., Wedler H., Wambutt R., Purnelle B.,
RA Goffeau A., Cadieu E., Dreano S., Gloux S., Lelaue V., Mottier S.,
RA Galibert F., Aves S.J., Xiang Z., Hunt C., Moore K., Hurst S.M.,
RA Lucas M., Rochet M., Gallardin C., Tallada V.A., Garzon A., Thode G.,
RA Dags R.R., Cruzado L., Jimenez J., Sanchez M., del Rey F., Benito J.,
RA Dominguez A., Revuelta J.L., Moreno S., Armstrong J., Forsberg S.L.,
RA Shpakovski G., Lowe T., McCombie W.R., Paulsen I., Potashkin J.,
RA Shpakovski L.V., Ussery D., Barrell B.G., Nurse P.;
RA "The genome sequence of Schizosaccharomyces pombe.";
RA Nature 415:871-880(2002).
RN [2]
RP FUNCTION, AND SUBCELLULAR LOCATION.
RP MEDLINE=21924702; PubMed=11927555;
RP Krien M.J.E., West R.R., John U.P., Koniaras K., McIntosh J.R.,
RP O'Connell M.J.;
RP "The fission yeast NIMA kinase Finlp is required for spindle function
RP and nuclear envelope integrity.";
RP EMBO J. 21:1713-1722(2002).
CC -!- FUNCTION: Promotes chromosome condensation and nuclear envelope
CC dynamics during mitosis. Activity appears at metaphase-anaphase
CC transition.
CC -!- SUBCELLULAR LOCATION: Spindle pole body.
CC -!- SIMILARITY: Belongs to the Ser/Thr family of protein kinases. NIMA
CC subfamily.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/
CC or send an email to license@isb-sib.ch)
CC
CC EMBL; Z98975; CAB11653.1; -
CC F01; T37970; T37970.
CC GeneDB SPombe; SPAC19E9.02; -
CC InterPro; IPR000719; Prot kinase.
CC InterPro; IPR008271; Ser thr pkin AS.

```

Corlieu A., Cadieu E., Briano S., Groux S., Lelaire V., Mottier S.,  
 Galibert F., Aves S.-J., Xiang Z., Hunt C., Moore K., Hurst S.M.,  
 Lucas M., Rochet M., Galliardin C., Tallada V.A., Garzon A., Thode G.,  
 Daga R.R., Cruzado L., Jimenez J., Sanchez M., del Rey F., Benito J.,  
 Dominguez A., Revuelta J.L., Moreno S., Armstrong J., Forsburg S.L.,  
 Cerruti L., Lowe T., McCombie W.R., Paulsen I., Potaashkin J.,  
 Shpakowski G.V., Ussery D., Barrell B.G., Nurse P.;  
 "The genome sequence of *Schizosaccharomyces pombe*.";  
 Nature 415:871-880(2002).  
 [2]  
 FUNCTION, AND SUBCELLULAR LOCATION.  
 MEDLINE=21924702; PubMed=11927555;  
 Krien M.J.E., West R.R., John U.P., Koniaras K., McIntosh J.R.,  
 O'Connell M.J.;  
 "The fission yeast NIMA kinase Finlp is required for spindle function  
 and nuclear envelope integrity.";  
 EMBO J. 21:1713-1722(2002).  
 !=- FUNCTION: Promotes chromosome condensation and nuclear envelope  
 dynamics during mitosis. Activity appears at metaphase-anaphase  
 transition.  
 !=- SUBCELLULAR LOCATION: Spindle pole body.  
 !=- SIMILARITY: Belongs to the Ser/Thr family of protein kinases. NIMA  
 subfamily.  
 -----  
 This SWISS-PROT entry is copyright. It is produced through a collaboration  
 between the Swiss Institute of Bioinformatics and the EMBL outstation  
 the European Bioinformatics Institute. There are no restrictions on its  
 use by non-profit institutions as long as its content is in no way  
 modified and this statement is not removed. Usage by and for commercial  
 entities requires a license agreement (see <http://www.isb-sib.ch/announce/>  
 or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 -----  
 EMBL; Z98975; CAB11653.1; -;  
 FIR; T37970; T37970.  
 GeneDB SPombe; SPAC19E9.02; -;  
 InterPro; IPR000719; Prot kinase.  
 InterPro; IPR008271; Ser thr pkin AS.  
 -----

```
DR InterPro; IPR002290; Ser_thr_kinase.
DR InterPro; IPR001245; Tyr_kinase.
DR Pfam; PF00069; kinase; 1.
DR PRINTS; PD00109; TYRKINASE.
DR ProDom; PD000001; Prot_kinase; 1.
DR SMART; SM00220; S_TKc_1.
DR PROSITE; PS00107; PROTEIN_KINASE_ATP; FALSE_NEG.
DR PROSITE; PS00111; PROTEIN_KINASE_DOM; 1.
DR PROSITE; PS0108; PROTEIN_KINASE_SF; 1.
KW Transferrase; Serine/threonine-protein kinase; ATP-binding; Mitosis;
KW Cell cycle; Cell division.
FT DOMAIN 4 281 PROTEIN KINASE.
FT NP_BIND 10 38 ATP (BY SIMILARITY).
FT BINDING 33 33 ATP (BY SIMILARITY).
FT ACT_SITE 151 151 BY SIMILARITY.
SQ SEQUENCE 722 AA; 82689 MW; 20C7E304DAD7C440 CRC64;

Query Match 17.9%; Score 448; DB 1; Length 722;
Best Local Similarity 26.8%; Pred. No. 3.3e-16;
Matches 121; Conservative 96; Mismatches 169; Indels 66; Gaps 14;

QY 28 RYVLOKLGSGPTVYLVSDKAKGEBELKVLKEISVGLNPNVQANLEAQLLSKLD 87
DB 3 KYKILECIGHGSGRTYKV--QRLKDG-ALLAQKEIHFQNTROEKQYIADENVILNLK 59
QY 88 HPAIVKPHASFVQDNFCI--ITEYCEGRDLDKIQEYKQAGKIPENQIIEWFTQLLIG 145
DB 60 HPNIVQCEELNRSQVNLWYECGDLANLQRYKEEKRTQEVLKFFQLLLA 119
QY 146 VDYMH-----ERRILRLKSKVFL-KNVLKIGDFGVSLMGCS 185
DB 120 LYRCHYGENAPACDSQWPREIHPKQSVLHRDIPANIFLDENNSVYKLGDFGLSLDNT 179
QY 186 CDLATLTITGPHYSKALHQGYDFTKSDIWSLACILYEMCCWNAFAGSNFLSVLKV 245
DB 180 RVFTQSVGVGPPYNSPEIRSSPYSAKSDVWALGCVIFEICMLTHPFGRSGYLEQRNIC 239
QY 246 EGTPTSLPERYPKELNAINESMLNKNPSLRPAIEILKIPYLDQQLNLMCRYSMTLED 305
DB 240 QGNLSCWDHYSDDVFLIRHCELVNSDLRPTTYQLLRSPILSDIRSKL--ESERVLEQ 297
QY 306 KNIDCOKEAAHINAMQKRIHLOTLRALSEVQVQWTPRMRRLKQAADEKARKLKIVE 365
DB 298 SDL-----LHKH-QHMLIQENDLQ--FREQLSARESELENVIAISRLAQREE 342
QY 366 EKYEENSKRMOELSRNFQOLSVDVLHKEHLKGMEEKEEQEGRSLSCSPQDEDERWQG 425
DB 343 ILBRELEKQURDMDAR-YQR-----HMQTVNSMQKMR-----VTSFVDHNSQPESS 388
QY 426 REE-----ESDEPTLENLPE---SQPTSPM 447
DB 389 TAEWFVDCITTEASQSPLLHPKLGISKPLQTL 420

RESULT 14
NEK7_MOUSE STANDARD; PRT; 302 AA.
AC Q9ES74;
DT 10-OCT-2003 (Rel. 42, Created)
DT 10-OCT-2003 (Rel. 42, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Serine/threonine-protein kinase Nek7 (EC 2.7.1.37) (Nima-related
DE protein kinase 7).
GN NEK7.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=20422672; PubMed=10964517;
RA Kandli M., Feige E., Chen A., Kilfin G., Motro B.;
RT "Isolation and characterization of two evolutionarily conserved murine

kinases (Nek6 and Nek7) related to the fungal mitotic regulator,
NIMA";
Genomics 68:187-196(2000).
[2]
SEQUENCE FROM N.A.
STRAIN=C57BL/6J, and NOD; TISSUE=Thymus, and Urinary bladder;
MEDLINE=22354683; PubMed=12466851;
RA Okazaki Y., Furuno M., Kasukawa T., Suzuki H., Yamanaka I., Kiyosawa H.,
RA Nikaide I., Osuto N., Saito R., Nogi A., Schonbach C., Gojibori T.,
RA Yagi K., Tomaru Y., Hasegawa Y., Nogi A., Hume D.A., Quackenbush J.,
RA Baldarelli R., Hill D.P., Bult C., Hume D.A., Quackenbush J.,
RA Schriml L.M., Kanapin A., Matsuda H., Batalov S., Corbani L.E., Cousins S.,
RA Blake J.A., Bradt D., Brusic V., Chothia C., Corbani L.E., Cousins S.,
RA Dalla E., Dragani T.A., Fletcher C.F., Forrest A., Frazer K.S.,
RA Gaasterland T., Gariboldi M., Gissi C., Godzik A., Gough J.,
RA Grimmond S., Gustincich S., Hirokawa N., Jackson I.J., Jarvis E.D.,
RA Kanai A., Kawai H., Kawasawa Y., Redziarski R.M., King B.L.,
RA Konagaya A., Kurochkin I.V., Lee Y., Lenhard B., Lyons P.A.,
RA Maglott D.R., Maltais L., Marchionni L., McKenzie L., Miki H.,
RA Nagashima T., Numata K., Okido T., Pavan W.J., Pertea G., Pesole G.,
RA Petrovsky N., Pillai R., Pontius J.U., Qi D., Ramachandran S.,
RA Ravasi T., Reed J.C., Reed D.J., Reid J., Ring B.Z., Ringwald M.,
RA Sandelin A., Schneider C., Semple C.A., Setou M., Shimada K.,
RA Sultana R., Takenaka Y., Taylor M.S., Teasdale R.D., Tomita M.,
RA Verardo R., Wagner L., Walestedt C., Wang Y., Watanabe Y., Wells C.,
RA Wilming L.G., Wyszynski B.A., Yanagisawa M., Yang I., Yang L.,
RA Yuan Z., Zavolan M., Zhu Y., Zimmer A., Carninci P., Hayatsu N.,
RA Hirozane-Kishikawa T., Konno H., Nakamura M., Sakazume N., Sato K.,
RA Shiraki T., Waki K., Kawai J., Aizawa K., Arakawa T., Fukuda S.,
RA Hara A., Hashizume W., Imotani K., Ishii Y., Itoh M., Kagawa I.,
RA Miyazaki A., Sakai K., Sasaki D., Shibata K., Shinagawa A.,
RA Yasunishi A., Yoshino M., Waterston R., Lander E.S., Rogers J.,
RA Birney E., Hayashizaki Y.,
RT "Analysis of the mouse transcriptome based on functional annotation of
RT 60,770 full-length cDNAs";
RL Nature 420:563-573(2002).
[3]
SEQUENCE FROM N.A.
STRAIN=Czech II;
MEDLINE=22388257; PubMed=12477932;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Ustin T.B., Toshiyuki S., Carninci P., Frange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullany S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahey J., Heiton E., Kettman M., Madan A., Rodriguez S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko I., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Myers R.M.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA Butterfield V.S.N., Krzywinski M.I., Skalska U., Smalls D.E.,
RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human
RT and mouse cDNA sequences";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
CC -!- CATALYTIC ACTIVITY: ATP + a protein = ADP + a phosphoprotein.
CC -!- SUBCELLULAR LOCATION: Cytoplasmic.
CC -!- SIMILARITY: Belongs to the Ser/Thr family of protein kinases. NIMA
CC subfamily.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@sib-sib.ch).
```







---

**This Page Blank (uspto)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: October 1, 2004, 11:13:40 ; Search time 41 Seconds

(without alignments)  
1130.837 Million cell updates/sec

Title: US-10-803-278-4

Perfect score: 2501

Sequence: 1 MLKQFAAKCVSGSSTAIITY.....LGVHGDNCNLISLDFWKNK 482

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283366 seqs, 96191526 residues

Total number of hits satisfying chosen parameters: 283366

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

PIR 78:\*\*

1: p1r1:\*\*

2: p1r2:\*\*

3: p1r3:\*\*

4: p1r4:\*\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	615	24.6	841	1 I78885	serine/threonine-s
2	600.5	24.0	792	2 JC7122	protein kinase (PC
3	598	23.9	774	2 S25284	protein kinase nek
4	495	19.8	445	2 G01452	NIMA-like protein
5	463.5	18.5	435	2 S23580	probable protein k
6	448	17.9	722	2 T37970	probable G2-specif
7	440.5	17.6	699	2 A43734	probable protein k
8	439	17.6	338	2 JC7838	Nek6 protein kinase
9	434.5	17.4	941	2 T49136	protein kinase-like
10	428.5	17.1	357	2 T29771	hypothetical prote
11	414	16.6	294	2 T21075	hypothetical prote
12	410.5	16.4	779	2 A57177	NIMA-like protein
13	409.5	16.4	431	2 T11854	protein kinase (BC
14	392.5	15.7	158	2 T29253	hypothetical prote
15	388	15.5	446	2 A40896	Ca2+/calmodulin-de
16	387	15.5	1233	2 T14157	serine/threonine p
17	382.5	15.3	1062	2 S46367	protein kinase CDC
18	376	15.0	848	2 T47986	serine/threonine-p
19	375	15.0	1231	2 T19532	serine/threonine pr
20	372.5	14.9	682	2 A44493	serum-inducible ki
21	370	14.8	200	2 B96587	hypothetical prote
22	368.5	14.7	447	2 B40896	Ca2+/calmodulin-de
23	362.5	14.5	795	2 JC4234	gene fused protein
24	362	14.5	651	2 A96591	NPK1-related prote
25	358	14.3	1206	2 T34021	protein kinase SK2
26	355.5	14.2	1895	2 T15681	hypothetical prote
27	354	14.2	603	2 S34130	serine/threonine-s
28	353.5	14.1	1097	2 F96538	hypothetical prote
29	345	13.8	1233	2 T30989	serine/threonine p

30 344 13.8 469 2 B84644  
31 344 13.8 608 2 G96575  
32 342.5 13.7 445 2 T50802  
33 342 13.7 705 2 A48144  
34 340 13.6 426 2 S71886  
35 340 13.6 461 2 T48222  
36 340 13.6 603 2 A54596  
37 339.5 13.6 836 2 B96716  
38 337 13.5 273 2 S11380  
39 337 13.5 603 2 A47545  
40 336.5 13.5 974 2 S15038  
41 335.5 13.4 421 2 T48202  
42 335.5 13.4 738 1 S51380  
43 335.5 13.4 897 2 S61137  
44 335 13.4 435 2 B84707  
45 333 13.3 394 2 JU0229

#### ALIGNMENTS

##### RESULT 1

I78885

serine/threonine-specific protein kinase (EC 2.7.1.1-) STK2 - human

C:Species: Homo sapiens (man)

C>Date: 10-Sep-1999 #sequence\_revision 10-Sep-1999 #text\_change 29-Sep-1999

C:Accession: I78885

R:Revedakou, E.N.; He, M.; Baptist, E.W.; Craven, R.J.; Cance, W.G.; Welcsh, P.L.; Simmor

Oncogene 9, 1977-1988, 1994

A:Title: Two novel human serine/threonine kinases with homologies to the cell cycle regul

A:Reference number: I58396; MUID:94268838; PMID:8208544

A:Accession: I78885

A>Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: mRNA

A:Residues: 1-841 <RES>

A:Cross-references: GB:L20321; NID:g348244; PIDN:AAA36658.1; PID:g348245

C:Genetics:

A:Gene: GDB:STK2

A:Cross-references: GDB:374125

A:Map position: 3p21.1-3p21.1

C:Superfamily: human serine/threonine-specific protein kinase STK2; protein kinase homolc

C:Keywords: phosphotransferase

F/4-261/Domain: protein kinase homology <KIN>

Query Match 24.6%; Score 615; DB 1; Length 841;  
Best Local Similarity 34.7%; Pred. No. 3.3e-17;  
Matches 157; Conservative 87; Mismatches 152; Indels 56; Gaps 17;  
QY 29 YVQQKLGSGSGFTYLVSDKAKRGELKYLKEISYVGLNPNETVQANLEAQLSLKLDH 88  
Db 6 YCLRVRVVGKSGYGEVTLV---KRRDQKQYVKKLNLNASSRRRAAEQAQLLSQLKH 62  
QY 89 PAIVKFEASFVEQDNFC-IITEYCEGDLDDKIQEYKQAGKIPPENQIIEWFQILLGVD 147  
Db 63 PNIVTYKESMEGGDGLLYVNGFCEGDLRYKLKE--QKGQLLENQVWFVQIAVALQ 120  
QY 148 YMHERRILHRLDKSNVFL-KNNLLKICDFGVSRLLMGSCDLATLTGTGTHYMSPEALKH 206  
Db 121 YLKEKHILHRLDKTQNVFLRTNIIKVDIGIARVLENHCDMASTLIGTPYKSPFLFSN 180  
QY 207 QGYDTKSDISLACILYEMCMONHAPAGSNFLSLVLKIVEGDTPLSPERPKEINAMES 266  
Db 181 KPNYKSDVWALGCVVEMATLKAFNAKDMNSIVYRIIEGKLFPAMPDYSPELAELIRT 240  
QY 267 MLNKNPRLPSAIEILKIPYLDEQLNLCRYSEMTLED--KNLDCQ-KEAAHIN--- 319  
Db 241 MLSKPEPERSVRSILRPVQIKQISFLEATKTKTSKNNIKNGDSOSKFPATVVSGEAE 300  
QY 320 AMQKRIHLQTLRA-----LSEVQKMTPRMRMLRKL--QAADKARKLKKIVVEKYEN 371  
Db 301 SNHEVTHPQLSSSGSQTYIMGEGKCLSQKPRASGLLKPSALKAKTQCQDL-----SN 355  
QY 372 SKNQELSRNFQQLSDVVLHEKTHLKGME-----KEEQPE-----GRLLSCSPQD 417

Db 356 TTELATISSVN-----IDIL-----PAKGRDSVSDGFVQENQPRYLDASNELGGI-CSISQ 405  
Qy 418 EDERWQGRREEDPTLENLPESQIPSMDL 449  
Db 406 VEEMXLQNTKSAQP--ENL-----IPWSSDI 432

RESULT 2  
JC7122  
protein kinase (EC 2.7.1.37) 2 - mouse (strain balb/c)  
N;Alternate names: serine (threonine) protein kinase  
C;Species: Mus musculus (house mouse)  
C;Date: 23-Mar-2001 #sequence\_revision 23-Mar-2001 #text\_change 31-Mar-2001  
C;Accession: JC7122  
R;Hayashi, K.; Igarashi, H.; Ogawa, M.; Sakaguchi, N.  
Biochem. Biophys. Res. Commun. 264, 449-456, 1999  
A;Title: Activity and substrate specificity of the murine Ser/Thr kinase STK2  
A;Reference number: JC7122; MUID:20001940; PMID:10529384  
A;Accession: JC7122  
A;Status: preliminary  
A;Molecule type: mRNA  
A;Residues: 1-792 <HAY>  
A;Cross-references: GB:AJ223071; NID:G4138208; PID:G4138209  
C;Genetics:  
C;Superfamily: human serine/threonine-specific protein kinase STK2; protein kinase homolog  
C;Keywords: phosphotransferase

Query Match 24.0%; Score 600.5; DB 2; Length 792;  
Best Local Similarity 31.5%; Pred. No. 1.2e-16;  
Matches 141; Conservative 90; Mismatches 143; Indels 73; Gaps 13;  
Qy 29 YVLOQKLGSGFTGVYLVSDKKAKEELKVLKESVGLNPNETVQANLEAQLSKLDH 88  
Db 6 YCMRVVGRSGSYGEVLV---KRRDQKQYVVKLNLRNASSRERRAARAEQAQLSOLKH 62  
Qy 89 PAIVKPHASVFEODNFC--IITEYCEGRDLDDKIQEYKQAGKIPFENQIIEWFTQLLQVD 147  
Db 63 PNIVTYKESWEGDGLLYVWGCCEGDLVYRKLKE--QKQLLPESQVVEFWQIAWALQ 120  
Qy 148 YMERRILRLDKSKNVFL--KNLLKIGDGVGRLLMGSCDLATTLTGTTPHYMSPEALKH 206  
Db 121 YLKEKHLHLRLDLATQNVFLRTNIIKVGDLGARVLNENHGDMASTLIGTPYNSPELFSN 180  
Qy 207 QGYDTKSDIWSLACILYEMCMHAFAGSNFLSVLKIIVEGDTSPSPERYPKELNAIMES 266  
Db 181 KPNYKSDVWALGCGVEMATLKHAFNAKMNLSLVRIIEGKLPMPKYSTELAEALIT 240  
Qy 267 MLNKNPSLRPSAIEILKIPYLDLQ-----QNLMSRYSEMVL-----ED 305  
Db 241 MLSRRPERSVRSILRPQYKHHISLFLBETAKTSKNNVKNCDSRAPFAVAVSRKEE 300  
Qy 306 KNLQ-----CKEAAHIIN-----AMQKRIHLQTLRALSEVQKTPRERMLRLKLA 352  
Db 301 SNTDVIHQPRSSGSAHWGDKLSCVDPIDGLPSRPSAGHGTHCKQDM-----NN 355  
Qy 353 ADEKARKLKKI-----VEEKYEENSKRMQELSRNFQQLSVSDLVHKEKHLKMGMEKEBQ 406  
Db 356 TGSCATISRNIDILPAERDSDANAGVQE-----SOPQHVAADAEVDVQ 401  
Qy 407 PEGRLSCSPQDEDERWQGRREESDEP 433  
Db 402 -----CSISQ--KERLQGNKTSQDQ 421

RESULT 3  
S25284  
protein kinase nek1 (EC 2.7.1.1) - mouse  
C;Species: Mus musculus (house mouse)  
C;Date: 17-Apr-1993 #sequence\_revision 17-Apr-1993 #text\_change 24-Sep-1999  
C;Accession: S25284  
R;Letwin, K.; Mizzen, L.; Motro, B.; Ben-David, Y.; Bernstein, A.; Pawson, T.

EMBO J. 11, 3521-3531, 1992  
A;Title: A mammalian dual specificity protein kinase, Nek1, is related to the NIMA cell cycle  
A;Reference number: S25284; MUID:93010942; PMID:1382974  
A;Accession: S25284  
A;Status: not compared with conceptual translation  
A;Molecule type: mRNA  
A;Residues: 1-774 <LET>  
A;Cross-references: GB:S45828; NID:G256854; PIDN:AB23529.1; PID:G256855  
C;Genetics:  
A;Gene: nek1  
C;Superfamily: unassigned Ser/Thr or Tyr-specific protein kinases; protein kinase homolog  
C;Keywords: ATP; phosphotransferase  
F;2-58/Domain: protein kinase homology <KIN>  
F;10-18/Region: protein kinase ATP-binding motif

Query Match 23.9%; Score 598; DB 2; Length 774;  
Best Local Similarity 28.5%; Pred. No. 1.4e-16;  
Matches 141; Conservative 94; Mismatches 132; Indels 128; Gaps 13;  
Qy 28 RYVLOQKLGSGFTGVYLVSDKKAKEELKVLKESVGLNPNETVQANLEAQLSKLD 87  
Db 3 KYVRLQKIGSGFGKAVLV--KSTEDGRHY-VIKEINISRYMSDKERQESRREAVLANMK 59  
Qy 88 HPALVKPHASVFEODNFC--IITEYCEGRDLDDKIQEYKQAGKIPFENQIIEWFTQLLQVD 147  
Db 60 HPNIVQYKESFEENGSLYIVMDYCEGDLFGRIN--AQKALFQEDQILDWFOVQICLALK 117  
Qy 148 YMERRILRLDKSKNVFL--KNLLKIGDGVGRLLMGSCDLATTLTGTTPHYMSPEALKH 206  
Db 118 HVDRKILHRDIKSNIFLTKDGTQVQLGDFGARVLNSTVELARTCIGTPYLSPEICEN 177  
Qy 207 QGYDTKSDIWSLACILYEMCMHAFAGSNFLSVLKIIVEGDTSPSPERYPKELNAIMES 266  
Db 178 KPNYKSDIWSLACILYELCTLKHAFNAKMNLSLVRIIEGKLPMPKYSTELAEALIT 237  
Qy 267 MLNKNPSLRPSAIEILKIPYLDLQNLN-----C----- 296  
Db 238 LFKRNPDRPSVNSILEKGFIAKRIEFLSPQIAEAEFCLTKLSKFGQPLPGKRPASGQ 297  
Qy 297 -----RYSEMTLEDKNDLCOKEAAHIINAMQKRIHLQTLR 331  
Db 298 GVSSFVPAQKITTPAAKYGVPLTYKYGDKKLEKXPPKHKQAHQI----- 344  
Qy 332 ALSEVQKTPRERMLRLKLAQNADEKARKLKIIVEEKYEENSKRMQELSRNFQQLSV 391  
Db 345 ---FVKXNNSGEERKKMSEEA--KGRLEFIEKEK-----KQDKQIRFLKAEQMK--R 391  
Qy 392 HEKHLKMGMEKEBQ-----EGRLS-----CSPQ----- 416  
Db 392 QEKQLERINRAREQGWNVLRAGSGEVKASFPFGIGGAVSPSPSPRGQVHYHAFDQ 451  
Qy 417 -----DEDERWQ 425  
Db 452 MQLRAEDNEARWKG 466

RESULT 4  
G01452  
NIMA-like protein kinase 1 - human  
C;Species: Homo sapiens (man)  
C;Date: 21-Dec-1996 #sequence\_revision 06-Jun-1997 #text\_change 24-Sep-1999  
C;Accession: G01452  
R;Liu, K.P.  
Submitted to the EMBL Data Library, June 1994  
A;Reference number: G07172  
A;Accession: G01452  
A;Status: preliminary; translated from GB/EMBL/DBJ  
A;Molecule type: mRNA  
A;Residues: 1-445 <LUX>  
A;Cross-references: EMBL:U11050; NID:G507874; PIDN:AAA19558.1; PID:G507875  
C;Genetics:  
A;Gene: NIK1  
C;Superfamily: unassigned Ser/Thr or Tyr-specific protein kinases; protein kinase homolog



A;Residues: 1-722 <BAD>  
A;Cross-references: EMBL:Z98975; PIDN: CAB11693.1; GSPDB: GNO0066; SPDB: SPAC19E9.02  
A;Experimental source: strain 972h; cosmid c19E9  
C;Genetics:  
A;Gene: SPDB: SPAC19E9.02  
A;Map position: 1  
A;Introns: 20/3, 28/3  
A;Superfamily: unassigned Ser/Thr or Tyr-specific protein kinases; protein kinase homolog  
C;Keywords: phosphotransferase; protein kinase

Query Match 17.9%; Score 448; DB 2; Length 722;  
Best Local Similarity 26.8%; Pred. No. 8e-11;  
Matches 121; Conservative 96; Mismatches 169; Indels 66; Gaps 14;

QY 28 RYVLOQKLGSGFGTVYLVSDKAKRGELKVLKEISVGLNPNETVQANLEAQLSKLD 87  
DB 3 KYKLEICGHGFGRIYKV--QRLKDG-ALLAQKEIHFGNITROEKYIADEVNILENLK 59  
QY 88 HPAIVKFAHFVEQDNFCI--ITEYCEGRDLDDKIQYQKAGKIPFENQIIEWFIQLLLG 145  
DB 60 HENIVQYCGEELNRSQAQVNLNLYMEYCGHGLANLIQRYKEEKKEFTQEVLFKFTQLLLA 119  
QY 146 VDMYH-----ERRILHRDLKSKNVFL-KNNLLKIGDFGVSRLMGS 195  
DB 120 LVRYHGENAPACDSQWPRETFHPKQSVLHRDIPANIFLDENNSVXLGDFGLSKLDNT 179  
QY 186 CDLATTLTGTPTMYSPFALKHQGYDTKSDIWSLACILYEMCMNHAFAGNFSILVKIV 245  
DB 180 RVFTQSYVGTPTMYSPFELIRSPYSKSDVWALGVIFEICMLTHPPEGRSYLELQRNIC 239  
QY 245 EGDTSLEPYPKELNATMESMLNKNPSLRPSAIEILKIPYLBOLQNLNLCRYSEMTELD 305  
DB 240 QGNLSWDHSDVDFVLRHCLVNSDLRPTTYQLLRSPILSDIRSKL--ESERVVLEQ 297  
QY 306 KNLDCQKAAHIINAMQKRIHLQTLRALSEVQKQTPRMRRLKQLQAADEKAKLAKIVE 365  
DB 298 SDL-----LHKK-HQMLIQLENDIQ--FREQLRSARESELENVIAKSLAQRE 342  
QY 366 EKYENSKRMQELSRNQQQLSDVVLHKEHKLKQWKEKEQPEGRGLSCSPQDDEERWQ 425  
DB 343 ILRRELEKQLRDMAR-QYR-----HMQTQVNSMQQR-----VTSVPVDHNEQPESS 388  
QY 426 REE-----ESDEPTLENLPE---SQPIPSM 447  
DB 389 TAEVFDCTIEASQSLAHIFKLIGSKPLQLT 420

RESULT 7  
A43734  
probable protein kinase nima (EC 2.7.1.1) - Emericella nidulans  
C;Species: Emericella nidulans, Aspergillus nidulans  
C;Date: 03-Mar-1993 #sequence\_revision 03-Mar-1993 #text\_change 24-Sep-1999  
C;Accession: A43734; S28786  
R;Osmani, S.A.; Pu, R.T.; Morris, N.R.  
Call 53, 237-244, 1988  
A;Title: Mitotic induction and maintenance by overexpression of a G2-specific gene that  
A;Reference number: A43734; MUID: 88194523; PMID: 3359487  
A;Accession: A43734  
A;Molecule type: DNA  
A;Residues: 1-699 <OSM>  
A;Cross-references: GB:M20249; NID: g168065; PIDN: AAA33316.1; PID: g168066  
C;Genetics:  
A;Gene: nima  
C;Superfamily: unassigned Ser/Thr or Tyr-specific protein kinases; protein kinase homolog  
C;Keywords: ATP; autophosphorylation; nucleus; phosphoprotein; phosphotransferase; serin  
P;9-295/Domain: protein kinase homology <KIN>  
P;17-25/Region: protein kinase ATP-binding motif

Query Match 17.6%; Score 440.5; DB 2; Length 699;  
Best Local Similarity 28.2%; Pred. No. 1.5e-10;  
Matches 133; Conservative 79; Mismatches 200; Indels 59; Gaps 14;

QY 26 ARRYVLOQKLGSGFGTVYLVSDKAKRGELKVL--KEISVGLNPNETVQANLEAQLL 83

DB 8 ADKVEVLEKICGSGFGII-----RKVRKSDGFLICRKEINYIMKSTKEREQLTASFNIL 62  
QY 84 SKLDHPAIVK-FHASFVE--QDNFCIITEYCEGRDLDDKIQYQKAGKIPFENQIIEWFI 140  
DB 63 SSLRHPNIVAYHREHLKASQDLY-LYMEYCGGGLSWIKNLKRTKRYABEDFVWRILS 121  
QY 141 QLLIGVDYMH-----ERRILHRDLKSKNVFL-KNNLLK 172  
DB 122 QLVATLYRCHYGTDPAEVGSNLLGPAPKPSGLKQKQAMTILHRDLKPNIFLGSNTVK 181  
QY 173 IGDGVSRLMGSCLATTITGTPHYNSPEALKHQGYDTKSDIWSLACILYEMCMNHAF 232  
DB 182 IGDGFLSK-LHSHDFASTVYGTPTMYSPICAETKYLRSDIWAVGCMYELCQREPPF 240  
QY 233 AGSFSLIVLKIYVGDTPSPERYPKELNATMESMLNKNPSLRPSAIEILKIPYL----- 287  
DB 241 NARTHIQLVQKIREGKFAPLPDFYSSSELKNVIAASCLRVNPDHDPDTATLINTPVIELMR 300  
QY 288 DEQLNL--MCRYSEMTELDKNDL-----CKEAAHIINAMQKRIHLQ-TLRALSEVQ 337  
DB 301 EVELNLSRAARKREBEATMQKADVECAFALKEKQKQIRSELENSIRREWEVKARLEID 360  
QY 338 KMTPRVRRLKQAADKARKLKKIVVEKYENSKRMQELSR-NFQQLSDVVLHKEH 396  
DB 361 RQVQNELDKLRPECEVQDRVAQVEKQRNANYREDASLRSSGHSSQWSSSSSDSF 420  
QY 397 LKGMEEKS---EQPEGRGLSCSPQDDEERWQGREESDEPTLENLPESQPI 444  
DB 421 PSSDIDLSLESSTNKAALKPKESRTPTRSKTVVDSFMDIQMAEPSPI 471

RESULT 8  
JC7838  
Nek6 protein kinase, NIMA histone H3 kinase homolog - human  
C;Species: Homo sapiens (man)  
C;Date: 09-Dec-2002 #sequence\_revision 09-Dec-2002 #text\_change 31-Mar-2003  
C;Accession: JC7838  
R;Hashimoto, Y.; Akita, H.; Hibino, M.; Kohri, K.; Nakanishi, M.  
Biochem. Biophys. Res. Commun. 293, 753-758, 2002  
A;Title: Identification and characterization of Nek6 protein kinase, a potential human hc  
A;Reference number: JC7838; MUID: 22050088; PMID: 12054534  
A;Accession: JC7838  
A;Molecule type: mRNA  
A;Residues: 1-338 <HAS>  
A;Cross-references: DDBJ: AB026289  
C;Comment: This protein, a novel member of NIMA (never in mitosis, gene A) related kinase  
playing a conserved central role in regulating mitotic chromosome condensation and assem  
C;Genetics:  
A;Gene: nek6  
A;Map position: 9q33-34

Query Match 17.6%; Score 439; DB 2; Length 338;  
Best Local Similarity 35.3%; Pred. No. 9.3e-11;  
Matches 95; Conservative 55; Mismatches 101; Indels 18; Gaps 7;

QY 20 YPKTLIAR-----RYVLOQKLGSGFGTVYLVSDKAKRGELKVLKEISVGLNPN 72  
DB 57 HPNTLSFRCSLADFPQIEKIGRGQFSEVYKATCLDRKT-----VALKKVQIFEMDAK 110  
QY 73 TVQANL-EAQLSKLDHPAIVKFAHFVEQDNFCIITEYCEGRDLDDKIQYQKAGKIP 131  
DB 111 ARQCVKEIGLLQKLNHPNIIKLDSEFIEDNELNIVLELADAGLSQMKIPFKQKRIIP 170  
QY 132 ENQIIEFIQLLGVDYMHERRILHRDLKSKNVFL-KNNLLKIGDFGVSRLMGSCLAT 190  
DB 171 ERTWKYFVQLCSAVEHMSRRVMHRDIKPNVITATGWLKGLDGLGRFSSSETTAH 230  
QY 191 TLCTGTPHYNSPEALKHQGYDTKSDIWSLACILYEMCMNHAFAGS--NFLSLVKIVSGD 248  
DB 231 SLVGTPTMYSPFELIRHNGNFKSDIWSLGLCLLYEMALQSPFYGDKMLFSLCQKLEOD 290  
QY 249 TPSLP-ERYPKELNATMESMLNKNPSLRP 276

```
Db      291 YPPLFGSHYSEKRELVSICDPDHPOR 319
||| ||| ||| :| | :| | :| | ||
RESULT 9
T49136
protein kinase-like protein - Arabidopsis thaliana
N;Alternate names: Protein.F26G5.150
C;Species: Arabidopsis thaliana (mousse-ear cress)
C;Date: 02-Jun-2000 #sequence_revision 02-Jun-2000 #text_change 02-Jun-2000
C;Accession: T49136
R;/D'Angelo, M.; Vezzi, A.; Modesto, D.; Pigazzi, M.; Valle, G.; Mewes, H.W.; Rudd, S.; I
submitted to the Protein Sequence Database, April 2000
A;Reference number: Z25017
A;Accession: T49136
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-941 <DNA>
A;Cross-references: EMBL:AL33814; GSPDB:GN00061; ATSP:F26G5.150
A;Experimental source: cultivar Columbia; BAC clone F26G5
C;Genetics:
A;Gene: ATSP:F26G5.150
A;Map position: 3
A;Introns: 33/2; 55/3; 77/3; 93/2; 108/3; 135/3; 150/1; 166/3; 194/1; 210/3; 218/3; 237/3;
Query Match 17.4%; Score 434.5; DB 2; Length 941;
Best Local Similarity 27.1%; Pred.No. 3 3e-10;
Matches 138; Conservative 78; Mismatches 172; Indels 121; Gaps 19;
QY 28 RYVLOKLGGSGFTGVLYSDKKAKRGEEUUKVLKEISVGELNPNETVQANLEAQLSKLD 87
:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||
Db 7 QYELMEQIGRGAFGAAILVHHKAERK---KYVLKKIRLARQTERCCRRSAHQEMSLIARVQ 63
QY 88 HPAlVFHFASFVDQNF-CIIITYCGSRDLDDKIQEYKQAQGIFFPENQIIEWFIOLLGV 146
:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||
Db 64 HPVIVFEFKAWBEKGVCVIITVGYCGGGMAEIMLK--KNGVVTFPEBKCKWFTQILLAV 121
QY 147 DYMHERRILHRDLKSKNFVL-KNNLLKIGDFGSRLMGSCDLATTLTGTPHYMSPEALK 205
:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||
Db 122 EVLSHNVYLHRDLKCSNIFLTQQDVLRGLDGFLAKTKAD-DLTSSVWGTFNYMCPELLA 180
QY 206 HQGYDTFKSIDIWSLACLILEYCMMNHAPAGSNFLSVLK-IVEGDTPSIPIRYPKELNAI 263
:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||
Db 181 DPIYGKSIDWSJGCCCIYEAAVRPAKFADFMDAGLIKSKSTHGQNVYKEEFVRSVK--- 236
QY 264 MESMLNKPNLRP-----SAIEIL-----KIPYLDEOLQNLMCRYSEMITLE 304
Db 237 MASEILKHPLYQVYBQRPTLSAAITPEKPINREGRRSMAESQNSNSEKONFYYS 296
QY 305 DKNL-----DCOKEAAHIINAMOKRIHLQTLRA-----LSEVQ 337
Db 297 DKNIRYVPSPNGKVKTETDSGFVDDIEDLVHQQSANG-NLQSVSATPGDHGILKPVH 355
QY 338 K-----MTPRERMRLKQAADEKARKLKCVIEEKYEENSKMOBLRSR-----N 382
Db 356 SQCRPDVIOFRPKTIrNi-----MMVLKEEKARENGSPMRNRPSRVPTQKN 406
QY 383 FOOLSv-----DVLHE-----KTHLKGMEF--KEEQPF 408
Db 407 VETPSKIPKLGIDIAHSKTNASTPIPPSKLASDSARTPGSFPPKHMFPVIDSPKLPKN 466
QY 409 GRLLSCSP--QDEREERWQGREEDSPTL 435
Db 467 DRISPSPA KHAEAEAMS VKRRQTPTPL 495
RESULT 10
T29771
hypothetical protein ZC581.1 - Caenorhabditis elegans
C;Species: Caenorhabditis elegans
C;Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 17-Mar-2000
C;Accession: T29771
R./Waterston, B.; Gattung, S.; Le, T.T.
```

Query Match 16.6%; Score 414; DB 2; Length 294;  
Best Local Similarity 34.9%; Pred. No. 7.6e-10;  
Matches 88; Conservative 56; Mismatches 90; Indels 18; Gaps 7;

QY 21 PKTLIARRVYLOQKLGSGFGTVLVSDKKARKE-----ELKV-LKEISVGLNPNETVQ 75  
DB 15 PDKLSLELFIEIKIGQSEVF-----RAQWTVDLHVALKIQVFEVMDQKARQ 66

QY 76 ANL-EAQLLSKLDHPAIVKTHASFTQDNFCITTEYCEGRDLDKIQEYKQAGKIPFNQ 134  
DB 67 DCLKEIDLLKQLNHNVNVIYASFDINNQLNIVLELAEGADMSRMKPKKGRLIPEKT 126

QY 135 IIEWFTOLLGVDMYHERRILHRDLKSKNVFLK-NLLKIGDGVSRLLMGSCDLATTT 193  
DB 127 IKWIFVOLARALAHMSKRIHHRDIKPNVFTGNGIVKGLDGLGRFFSSKTTAAHSIV 186

QY 194 GTPHYSPEALKHGGYDTSKDIWSLACILYEMCMNHAFAGS--NPLSIVLKIVEGDTPS 251  
DB 187 GTPVYSPERIQESGYNFKSDLWSTGCLLYEMAALQSPFYGDKNLYSLCKIENCEYPP 246

QY 252 LP-ERIPKELNA 262  
DB 247 LPADIYSTQVSA 258

RESULT 12  
A57177  
NIMA-like protein kinase - Neurospora crassa  
C:Species: Neurospora crassa  
C>Date: 10-Nov-1995 #sequence\_revision 10-Nov-1995 #text\_change 11-Aug-2003  
C:Accession: A57177  
R:Pu, R.T.; Xu, G.; Wu, L.; Vierula, J.; O'Donnell, K.; Ye, X.S.; Osmani, S.A.  
J. Biol. Chem. 270, 18110-18116, 1995  
A:Title: Isolation of a functional homolog of the cell cycle-specific NIMA protein kinase  
A:Reference number: A57177; MUID:95355415; PMID:7629122  
A:Status: preliminary; not compared with conceptual translation  
A:Molecule type: DNA  
A:Residues: 1-779 <PUA>  
A:Cross-references: GB:L42573; NID:G1040682; PIDN:AAA80145.1; PID:G1040683  
C:Superfamily: unassigned Ser/Thr or Tyr-specific protein kinases; protein kinase homolog  
C:Keywords: ATP  
F:5-290/Domain: protein kinase homology <KIN>  
F:13-21/Region: protein kinase ATP-binding motif

Query Match 16.4%; Score 410.5; DB 2; Length 779;  
Best Local Similarity 31.2%; Pred. No. 2.4e-09;  
Matches 121; Conservative 65; Mismatches 151; Indels 51; Gaps 12;

QY 28 RYVLOQKLGSGFGTVLVSDKKARKEELKVL--KEISVGLNPNETVQANLEAQLLSK 85  
DB 6 KYELLEKTGHSGFII-----RKVRRKADGMILCKEISYLNKMSQKEREQLHAEFSLST 60

QY 86 LDHPAIV-KTHASVE-QDNFCITTEYCEGRDLDKIQEYKQAGKIPFNQIIEWFIQLL 143  
DB 61 LRHPNIVGYRREHLKATQDLHLYMEVCGDGLGRVIRNLKKNQYAESEFWSIFSQVL 120

QY 144 LGVDYME-----ERRILHRDLKSKNVFL-KNLLKIGDGF 176  
DB 121 TALYRCHGVDPPEVGKTVLGLSTARPKPSPGGMTILHRDLKPNVFLGEDNSVKLGDF 180

QY 177 GVSRLMGSCDLATTTCTGTHYMSPEALKHGGYDTSKDIWSLACILYEMCMNHAFAGSN 236  
DB 181 GLSK-VMOSHDPASTVGTPTPYMSPETCAARKYTKSDIWSLGLGIIYELCARPPFNKT 239

QY 237 FLSIVLKIVEGDTSPERPYPKELNAMESMLNKNPSLRPSAIELKIPYLDQQLNMC 296  
DB 240 HVQLVQKIKEGKIAPLPSVYGELFATIKDCLRVNPPRPTATLLNLP-----IVRLMR 294

QY 297 RYSENTLEDKLDCKEAAHIINAVQKEI-HQLTLRALSEVQKMTPRRMLRLKLOADE 355  
DB 295 KEKEVVEFSRTLTKEE-----TLNKAIRELDSKLSALETEKSSIRAEIDASLRREWEV 348

QY 356 KAR-KLKKIVEEKYEE-NSKRMQELRSR 381  
DB 349 KARLEIDRLVAQBIESLQKQFEQVQAR 376

RESULT 13  
T11854  
protein kinase (EC 2.7.1.1) - Trypanosoma brucei  
C:Species: Trypanosoma brucei  
C>Date: 16-Jul-1999 #sequence\_revision 16-Jul-1999 #text\_change 11-May-2000  
C:Accession: T11854  
R:Gale, M.J.; Parsons, M.  
Mol. Biochem. Parasitol. 59, 111-122, 1993  
A:Title: Trypanosoma brucei gene family encoding protein kinases with catalytic domains  
A:Reference number: 217363; MUID:93295429; PMID:8515773  
A:Accession: T11854  
A:Status: preliminary; translated from GB/EMBL/DBDJ  
A:Molecule type: mRNA  
A:Residues: 1-431 <GAL>  
A:Cross-references: EMBL:L03778; NID:G162169; PID:G162170  
C:Genetics:  
A:Note: nrka  
C:Superfamily: unassigned Ser/Thr or Tyr-specific protein kinases; protein kinase homolog  
C:Keywords: ATP, phosphotransferase

Query Match 16.4%; Score 409.5; DB 2; Length 431;  
Best Local Similarity 31.5%; Pred. No. 1.6e-09;  
Matches 85; Conservative 72; Mismatches 104; Indels 9; Gaps 5;

QY 28 RYVLOQKLGSGFGTVLVSDKKARKEELKVLKEISVGLNPNETVQANLEAQLLSKLD 87  
DB 19 KLYNKGVIGLSYGEAYVA---ESVEDGSLCAKVMDSLXMSQDKRYAQSEIKCLANCN 75

QY 88 HPAIVKTHASFTQDNFCITTEYCEGRDLDKIQEYKQAG--KIFPENQIIEWFIQLLG 145  
DB 76 HPNIIRYIEDHEENDRLIIVMEFADSGNLDEQI-KLRGSGDARYFQBHEALFLQLJCLA 134

QY 146 VDYMERILHRDLKSKNVFL-KNLLKIGDGVSRLLMGSCD--LATTITGTGTHYMSPE 202  
DB 135 LDYHSHKOLHRDIKSNVLTSTGLVKLGDGFGSHQYEDTVSGVASTPGCTPYVLAPE 194

QY 203 ALKHGGYDTSKDIWSLACILYEMCMNHAFAGSNFLSIVLKIVEGDTSPERPYPKELNA 262  
DB 195 LWNKRYNKADVWSLGVLLYEMGMKPPFASNLKGLMSKVLAGTYAPLPDSFSSEFKR 254

QY 263 IMESMLNKNPSLRPSAIELKIPYLDQQL 292  
DB 255 VVDGILVADPNDRPSVREIFQIPYINKGLK 284

RESULT 14  
T29253  
hypothetical protein B0496.3 - Caenorhabditis elegans  
C:Species: Caenorhabditis elegans  
C>Date: 15-Oct-1999 #sequence\_revision 15-Oct-1999 #text\_change 15-Oct-1999  
C:Accession: T29253  
R:Murray, J.; Le, T.T.  
submitted to the EMBL Data Library, May 1996  
A:Description: The sequence of C. elegans cosmid B0496.  
A:Reference number: Z20596  
A:Status: preliminary; translated from GB/EMBL/DBDJ  
A:Molecule type: DNA  
A:Residues: 1-1558 <MUR>  
A:Cross-references: EMBL:U58749; PIDN:AAB00636.1; GSPDB:GN00022; CESP:B0496.3  
A:Experimental source: strain Bristol N2; clone B0496  
C:Genetics:  
A:Gene: CESP:B0496.3  
A:Map position: 4  
A:Introns: 55/1; 100/3; 141/3; 226/3; 298/1; 468/3; 502/2; 609/3; 656/3; 795/3; 843/2; 90

Query Match 15.7%; Score 392.5; DB 2; Length 1558;





**This Page Blank (uspio)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 1, 2004, 11:12:20 ; Search time 121 Seconds  
(without alignments)

1256.858 Million cell updates/sec

Title: US-10-803-278-4

Perfect score: 2501

Sequence: 1 MLKQFAAKCVSGSTAISTY.....LGVHGCNLSLDEWYKNEK 482

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1017041 seqs, 315518202 residues

Total number of hits satisfying chosen parameters: 1017041

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

SPTREMBL\_25.\*  
1: sp\_archaea.\*  
2: sp\_bacteria.\*  
3: sp\_fungi.\*  
4: sp\_human.\*  
5: sp\_invertebrate.\*  
6: sp\_mammal.\*  
7: sp\_mmc.\*  
8: sp\_organelle.\*  
9: sp\_phage.\*  
10: sp\_plant.\*  
11: sp\_rodent.\*  
12: sp\_virus.\*  
13: sp\_vertebrate.\*  
14: sp\_unclassified.\*  
15: sp\_rvirus.\*  
16: sp\_bacteriap.\*  
17: sp\_archaeap.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match %	Length	DB	ID	Description
1	2501	100.0	482	4	Q8TBY1	Q8TBY1 homo sapien
2	2408	96.3	470	4	Q8NG65	Q8NG65 homo sapien
3	2408	96.3	645	4	Q8NG66	Q8NG66 homo sapien
4	2310	92.4	637	6	Q8WNU8	Q8WNU8 macaca fasc
5	1757.5	70.3	628	11	Q8C004	Q8C004 mus musculus
6	1752.5	70.1	438	11	Q8BL56	Q8BL56 mus musculus
7	1596.5	63.8	365	11	Q8BW62	Q8BW62 mus musculus
8	1480	59.2	463	4	Q8H5F4	Q8H5F4 homo sapien
9	1396	55.8	395	6	Q8WNT6	Q8WNT6 macaca fasc
10	593.5	23.7	849	13	Q7T299	Q7T299 brachydanio
11	589.5	23.6	424	11	Q8CD72	Q8CD72 mus musculus
12	586.5	23.5	614	11	Q8CCJ0	Q8CCJ0 mus musculus
13	586.5	23.5	627	11	Q7TSC3	Q7TSC3 mus musculus
14	585.5	23.4	489	4	Q8J023	Q8J023 homo sapien
15	577	23.1	302	11	Q8BSB6	Q8BSB6 mus musculus
16	575.5	23.0	509	11	Q99K72	Q99K72 mus musculus

17	558	22.3	336	11	Q8C6N6	Q8C6N6 mus musculus
18	546.5	21.9	416	10	Q9LTP35	Q9LTP35 arabidopsis
19	543.5	21.7	1123	5	Q86I06	Q86I06 dictyosteli
20	514	20.6	943	10	Q84CU5	Q84CU5 cryza sativ
21	511	20.4	606	10	Q9CAU7	Q9CAU7 arabidopsis
22	509	20.4	621	10	Q8SA64	Q8SA64 populus x c
23	504	20.2	555	5	Q8N9C3	Q8N9C3 leishmania
24	499.5	20.0	555	10	Q8RXT4	Q8RXT4 arabidopsis
25	499	20.0	568	10	Q8RX66	Q8RX66 arabidopsis
26	495	19.8	491	4	Q86XH2	Q86XH2 homo sapien
27	495	19.8	609	10	Q947T1	Q947T1 lycopersico
28	485	19.4	442	13	Q9W622	Q9W622 xenopus lae
29	485	19.4	442	13	Q7ZYE3	Q7ZYE3 xenopus lae
30	476	19.0	384	4	Q96QN9	Q96QN9 homo sapien
31	475.5	19.0	841	5	Q9VC32	Q9VC32 dirosophila
32	472.5	18.9	440	13	Q7ZUN2	Q7ZUN2 brachydanio
33	468	18.7	389	13	Q9W623	Q9W623 xenopus lae
34	465.5	18.6	299	5	Q8T755	Q8T755 xenopus lae
35	459	18.4	443	11	Q91Z18	Q91Z18 mus musculus
36	455.5	18.2	326	4	Q7Z634	Q7Z634 homo sapien
37	451	18.0	1057	5	Q815D5	Q815D5 plasmodium
38	450	18.0	443	11	Q921N9	Q921N9 mus musculus
39	450	18.0	579	5	Q95XQ3	Q95XQ3 caenorhabdi
40	445	17.8	310	13	Q7ZSX4	Q7ZSX4 xenopus lae
41	437.5	17.5	393	11	Q91XQ1	Q91XQ1 rattus norv
42	436.5	17.5	302	5	Q19530	Q19530 caenorhabdi
43	435	17.4	366	11	Q8BF64	Q8BF64 mus musculus
44	434.5	17.4	356	5	Q8N0P1	Q8N0P1 paramecium
45	434.5	17.4	941	10	Q9LXP3	Q9LXP3 arabidopsis

## ALIGNMENTS

### RESULT 1

Q8TBY1 ID Q8TBY1 PRELIMINARY; PRT; 482 AA.  
AC Q8TBY1  
DT 01-JUN-2002 (TrEMBLrel. 21, Created)  
DT 01-JUN-2002 (TrEMBLrel. 21, Last sequence update)  
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)  
DE Hypothetical protein.  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
OX NCBI\_TaxID=9606;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC TISSUE=Testis;  
RA Strausberg R.;  
RL Submitted (APR-2002) to the EMBL/GenBank/DBJ databases.  
CC -!- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.  
DR EXBL: BC028587; AAH28587.1; -.  
DR GO: GO:0005524; F:ATP binding; IEA.  
DR GO: GO:0004674; F:Protein serine/threonine kinase activity; IEA.  
DR GO: GO:0004713; F:Protein-tyrosine kinase activity; IEA.  
DR GO: GO:0016740; F:Transferase activity; IEA.  
DR GO: GO:0006458; P:Protein amino acid phosphorylation; IEA.  
DR InterPro: IPR000715; Prot\_kinase.  
DR InterPro: IPR002290; Ser\_thr\_kinase.  
DR InterPro: IPR006271; Ser\_thr\_pkin\_AS.  
DR InterPro: IPR001245; Tyr\_pkinase.  
DR Pfam: PF00069; pkinase; 1.  
DR PRINTS: PR00109; TYRKINASE.  
DR ProDom: PD000001; Prot\_kinase; 1.  
DR SMART: SM00220; S\_TKG\_1.  
DR PROSITE: PS00011; PROTEIN\_KINASE\_DOM; 1.  
DR PROSITE: PS00108; PROTEIN\_KINASE\_ST; 1.  
DR Hypothetical protein; ATP-binding; Kinase;  
KW Serine/threonine-protein kinase; Transferase.  
SQ SEQUENCE 482 AA; 55519 MW; FA6D89550C0223A5 CRC64;

Query Match 100.0%; Score 2501; DB 4; Length 482;

Best Local Similarity 100.0%; Pred. No. 2.3e-160;  
Matches 482; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MLKFOBAKCVSGSTAIPTKTLIARRVYVLOQKLGSGSGTGYLVSDKAKRGEELKVL 60  
DB 1 MLKFOBAKCVSGSTAIPTKTLIARRVYVLOQKLGSGSGTGYLVSDKAKRGEELKVL 60

QY 61 KEISVGLNPNETVQANLEAQLSKLDHPAIVKFAHAFVEQDNFCIIITEYCEGRDLDDKI 120  
DB 61 KEISVGLNPNETVQANLEAQLSKLDHPAIVKFAHAFVEQDNFCIIITEYCEGRDLDDKI 120

QY 121 QYKQAGKIFPENQIIIEWFIQLLGVDMYHERRILHRDLKSNVFLKNLLKIGDFGVS 180  
DB 121 QYKQAGKIFPENQIIIEWFIQLLGVDMYHERRILHRDLKSNVFLKNLLKIGDFGVS 180

QY 181 LLMGSCDLATTLTGTPHYMSPEALKHQGYDTKSDIWSLACILYEMCCMNHAFAGSNFLSI 240  
DB 181 LLMGSCDLATTLTGTPHYMSPEALKHQGYDTKSDIWSLACILYEMCCMNHAFAGSNFLSI 240

QY 241 VLKIVEGDTSLPERYPKELNAMESMLNKNPSLRPSAIEILKIPLYDEOLNLMCRYSE 300  
DB 241 VLKIVEGDTSLPERYPKELNAMESMLNKNPSLRPSAIEILKIPLYDEOLNLMCRYSE 300

QY 301 MTLEDKNDLCOKEAAHIINAMQKRIHLQTLRALSEVQKMTPRERMLRLKLADEKARKL 360  
DB 301 MTLEDKNDLCOKEAAHIINAMQKRIHLQTLRALSEVQKMTPRERMLRLKLADEKARKL 360

QY 361 KKIIVEEYKNSKMOELSRNFQOLSDVVLHEKTHLKGMEKEBQPEGLSCSPODEDE 420  
DB 361 KKIIVEEYKNSKMOELSRNFQOLSDVVLHEKTHLKGMEKEBQPEGLSCSPODEDE 420

QY 421 ERWQGREESDEPTLENLPESQIPSPMDLHELESIVEDATSDLYGDCNLSLDEYWN 480  
DB 421 ERWQGREESDEPTLENLPESQIPSPMDLHELESIVEDATSDLYGDCNLSLDEYWN 480

QY 481 EK 482  
DB 481 EK 482

## RESULT 2

Q8NG65 PRELIMINARY; PRT; 470 AA.

ID Q8NG65  
AC Q8NG65  
DT 01-OCT-2002 (TRENBLrel. 22, Created)  
DT 01-OCT-2002 (TRENBLrel. 22, Last sequence update)  
DT 01-OCT-2003 (TRENBLrel. 25, Last annotation update)  
DE NIMA-related kinase 11S.  
GN NEK11S.  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
OX NCBI\_TaxID=9606;  
RN [1]  
RP SEQUENCE FROM N.A.

RA Noguchi K., Fukazawa H., Murakami Y., Uehara Y.;  
RT "Nek11, a new member of the NIMA family of kinases, involved in DNA  
RT replication and genotoxic stress responses.";  
RL Submitted (SEP-2001) to the EMBL/GenBank/DBJ databases.  
DR EMBL; AB071997; BAC06331.1; -.  
DR GO; GO:0005524; F:ATP binding; IEA.  
DR GO; GO:0004674; F:protein serine/threonine kinase activity; IEA.  
DR GO; GO:0004713; F:protein-tyrosine kinase activity; IEA.  
DR GO; GO:0016740; P:transferase activity; IEA.  
DR GO; GO:0006468; P:protein amino acid phosphorylation; IEA.  
DR InterPro; IPR000719; Prot kinase.  
DR InterPro; IPR002290; Ser\_Thr\_kinase.  
DR InterPro; IPR008271; Ser\_Thr\_pkinase.  
DR Pfam; PF00069; pkinase; 1.  
DR PRINTS; PR00109; TYRKINASE.  
DR ProDom; PD000001; Prot kinase; 1.  
DR SMART; SM00220; S\_TKc; 1.

DR SMART; SM00219; TyrKc; 1.  
DR PROSITE; PS00011; PROTEIN KINASE DOM; 1.  
DR PROSITE; PS00108; PROTEIN\_KINASE\_ST; 1.  
KW ATP-binding; Kinase; Transferase.  
SQ SEQUENCE 470 AA; 54006 MW; F52B82ED2B096FFB CRC64;

Query Match 96.3%; Score 2408; DB 4; Length 470;  
Best Local Similarity 100.0%; Pred. No. 4.1e-154;  
Matches 466; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MLKFOBAKCVSGSTAIPTKTLIARRVYVLOQKLGSGSGTGYLVSDKAKRGEELKVL 60  
DB 1 MLKFOBAKCVSGSTAIPTKTLIARRVYVLOQKLGSGSGTGYLVSDKAKRGEELKVL 60

QY 61 KEISVGLNPNETVQANLEAQLSKLDHPAIVKFAHAFVEQDNFCIIITEYCEGRDLDDKI 120  
DB 61 KEISVGLNPNETVQANLEAQLSKLDHPAIVKFAHAFVEQDNFCIIITEYCEGRDLDDKI 120

QY 121 QYKQAGKIFPENQIIIEWFIQLLGVDMYHERRILHRDLKSNVFLKNLLKIGDFGVS 180  
DB 121 QYKQAGKIFPENQIIIEWFIQLLGVDMYHERRILHRDLKSNVFLKNLLKIGDFGVS 180

QY 181 LLMGSCDLATTLTGTPHYMSPEALKHQGYDTKSDIWSLACILYEMCCMNHAFAGSNFLSI 240  
DB 181 LLMGSCDLATTLTGTPHYMSPEALKHQGYDTKSDIWSLACILYEMCCMNHAFAGSNFLSI 240

QY 241 VLKIVEGDTSLPERYPKELNAMESMLNKNPSLRPSAIEILKIPLYDEOLNLMCRYSE 300  
DB 241 VLKIVEGDTSLPERYPKELNAMESMLNKNPSLRPSAIEILKIPLYDEOLNLMCRYSE 300

QY 301 MTLEDKNDLCOKEAAHIINAMQKRIHLQTLRALSEVQKMTPRERMLRLKLADEKARKL 360  
DB 301 MTLEDKNDLCOKEAAHIINAMQKRIHLQTLRALSEVQKMTPRERMLRLKLADEKARKL 360

QY 361 KKIIVEEYKNSKMOELSRNFQOLSDVVLHEKTHLKGMEKEBQPEGLSCSPODEDE 420  
DB 361 KKIIVEEYKNSKMOELSRNFQOLSDVVLHEKTHLKGMEKEBQPEGLSCSPODEDE 420

QY 421 ERWQGREESDEPTLENLPESQIPSPMDLHELESIVEDATSDLYGDCNLSLDEYWN 466  
DB 421 ERWQGREESDEPTLENLPESQIPSPMDLHELESIVEDATSDLYGDCNLSLDEYWN 466

## RESULT 3

Q8NG66 PRELIMINARY; PRT; 645 AA.

ID Q8NG66  
AC Q8NG66  
DT 01-OCT-2002 (TRENBLrel. 22, Created)  
DT 01-OCT-2002 (TRENBLrel. 22, Last sequence update)  
DT 01-OCT-2003 (TRENBLrel. 25, Last annotation update)  
DE NIMA-related kinase 11L.  
GN NEK11L.  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
OX NCBI\_TaxID=9606;  
RN [1]  
RP SEQUENCE FROM N.A.

RA Noguchi K., Fukazawa H., Murakami Y., Uehara Y.;  
RT "Nek11, a new member of the NIMA family of kinases, involved in DNA  
RT replication and genotoxic stress responses.";  
RL Submitted (SEP-2001) to the EMBL/GenBank/DBJ databases.  
DR EMBL; AB071996; BAC06350.1; -.  
DR GO; GO:0005524; F:ATP binding; IEA.  
DR GO; GO:0004674; F:protein serine/threonine kinase activity; IEA.  
DR GO; GO:0004713; F:protein-tyrosine kinase activity; IEA.  
DR GO; GO:0016740; P:transferase activity; IEA.  
DR GO; GO:0006468; P:protein amino acid phosphorylation; IEA.  
DR InterPro; IPR000719; Prot kinase.  
DR InterPro; IPR002290; Ser\_Thr\_pkinase.  
DR InterPro; IPR008271; Ser\_Thr\_pkinase.  
DR SMART; SM00220; S\_TKc; 1.

DR Pfam, PF00069; pkinase; 1.  
DR PRINTS; PR00109; TYRKINASE.  
DR ProDom, PD000001; Prot Kinase; 1.  
DR SMART; SM00220; S\_TK; 1.  
DR SMART; SM00219; TYK; 1.  
DR PROSITE; PS0011; PROTEIN KINASE DOM; 1.  
DR PROSITE; PS00108; PROTEIN KINASE ST; 1.  
KW ATP-binding; Kinase; Transferase.  
SQ SEQUENCE 645 AA; 74162 MW; FE957805CSFFB059 CRC64;

Query Match 96.3%; Score 2408; DB 4; Length 645;  
Best Local Similarity 100.0%; Pred. No. 6e-154;  
Matches 466; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MLKFOEAAKCVSGSTAISTYPTKTLIARRVYVLOQLGSGSGFTVYVSDKKARGEELKVL 60  
DB 1 MLKFOEAAKCVSGSTAISTYPTKTLIARRVYVLOQLGSGSGFTVYVSDKKARGEELKVL 60  
QY 61 KEISVGEINPNETVQANLEAQLLSKLDHPAIVKPHASFEVDNFCIITEYCEGRDLDDKI 120  
DB 61 KEISVGEINPNETVQANLEAQLLSKLDHPAIVKPHASFEVDNFCIITEYCEGRDLDDKI 120  
QY 121 QEYKQAGKIPFENQIIEWFIQLLGVDMYHERRILHRDLKSNVFLKNNLLKIGDFGVSR 180  
DB 121 QEYKQAGKIPFENQIIEWFIQLLGVDMYHERRILHRDLKSNVFLKNNLLKIGDFGVSR 180  
QY 181 LMGSCDLATLTGTTHYMSPEALKHQGYDTKSDIWSLACILYEMCCMNHAFAGSNFLSI 240  
DB 181 LMGSCDLATLTGTTHYMSPEALKHQGYDTKSDIWSLACILYEMCCMNHAFAGSNFLSI 240  
QY 241 VLKIVEGTPSLPERYPKELNAIVESMLNKNPSLRPSAIEILKIPYLDLQNLMLCRYS 300  
DB 241 VLKIVEGTPSLPERYPKELNAIVESMLNKNPSLRPSAIEILKIPYLDLQNLMLCRYS 300  
QY 301 MTEDEKNDLCKEAAHIINAMQKRIHLOTLRALSEVQKMTPRERMLRKLQAADKARKL 360  
DB 301 MTEDEKNDLCKEAAHIINAMQKRIHLOTLRALSEVQKMTPRERMLRKLQAADKARKL 360  
QY 361 KKIIVEEKEENSKMQLSRNFQOLSVDVLEKTHLKGMEKEEQEGRGLSCSPQDEDE 420  
DB 361 KKIIVEEKEENSKMQLSRNFQOLSVDVLEKTHLKGMEKEEQEGRGLSCSPQDEDE 420  
QY 421 ERWQGREESDEPTLENLPESQIPSMDLHELESIVEDATSDLYH 466  
DB 421 ERWQGREESDEPTLENLPESQIPSMDLHELESIVEDATSDLYH 466

RESULT 4  
Q8WN08 PRELIMINARY; PRT; 637 AA.  
ID Q8WN08  
AC Q8WN08;  
DT 01-MAR-2002 (TremBLrel. 20, Created)  
DT 01-MAR-2002 (TremBLrel. 20, Last sequence update)  
DT 01-OCT-2003 (TremBLrel. 25, Last annotation update)  
DE mRNA, similar to human hypothetical protein FLJ23495, complete cds.  
OS Macaca fascicularis (Crab eating macaque) (Cynomolgus monkey).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Cercopithecoidea;  
OC Cercopithecoidea; Macaca.  
OX NCBI\_TaxID=9541;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC TISSUE=Testis;  
RA Hashimoto K., Osada N., Hida M., Kuesda J., Tanuma R., Hirai M.,  
RA Terao K., Suzuki Y., Sugano S.;  
RT "Isolation of novel full-length cDNA clones from macaque testis cDNA libraries."  
RL Submitted (JUL-2001) to the EMBL/GenBank/DBJ databases.  
CC -!- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.  
KW EMBL; AB064997; BAB83539.1; -.  
DR GO; GO:0005524; F:ATP binding; IEA.  
DR GO; GO:0004674; F:protein serine/threonine kinase activity; IEA.

DR GO; GO:0004713; F:protein-tyrosine kinase activity; IEA.  
DR GO; GO:0016740; P:transferase activity; IEA.  
DR GO; GO:0006468; P:protein amino acid phosphorylation; IEA.  
DR InterPro; IPR000719; Prot\_Kinase.  
DR InterPro; IPR002290; Ser\_thr\_kinase.  
DR InterPro; IPR008271; Ser\_thr\_kinase.  
DR InterPro; IPR001245; Tyr\_kinase.  
DR Pfam; PF00069; pkinase; 1.  
DR PRINTS; PR00109; TYRKINASE.  
DR ProDom; PD000001; Prot\_kinase; 1.  
DR SMART; SM00220; S\_TK; 1.  
DR PROSITE; PS0011; PROTEIN KINASE DOM; 1.  
DR PROSITE; PS00108; PROTEIN KINASE ST; 1.  
KW ATP-binding; Kinase; Serine/threonine-protein kinase; Transferase.  
SQ SEQUENCE 637 AA; 73337 MW; FA3A5C3CEEAD16F CRC64;

Query Match 92.4%; Score 2310; DB 6; Length 637;  
Best Local Similarity 95.9%; Pred. No. 2.3e-147;  
Matches 447; Conservative 9; Mismatches 10; Indels 0; Gaps 0;

QY 1 MLKFOEAAKCVSGSTAISTYPTKTLIARRVYVLOQLGSGSGFTVYVSDKKARGEELKVL 60  
DB 1 MLKFOEAAKCVSGSTAISTYPTKTLIARRVYVLOQLGSGSGFTVYVSDKKARGEELKVL 60  
QY 61 KEISVGEINPNETVQANLEAQLLSKLDHPAIVKPHASFEVDNFCIITEYCEGRDLDDKI 120  
DB 61 KEISVGEINPNETVQANLEAQLLSKLDHPAIVKPHASFEVDNFCIITEYCEGRDLDDKI 120  
QY 121 QEYKQAGKIPFENQIIEWFIQLLGVDMYHERRILHRDLKSNVFLKNNLLKIGDFGVSR 180  
DB 121 QEYKQAGKIPFENQIIEWFIQLLGVDMYHERRILHRDLKSNVFLKNNLLKIGDFGVSR 180  
QY 181 LMGSCDLATLTGTTHYMSPEALKHQGYDTKSDIWSLACILYEMCCMNHAFAGSNFLSI 240  
DB 181 LMGSCDLATLTGTTHYMSPEALKHQGYDTKSDIWSLACILYEMCCMNHAFAGSNFLSI 240  
QY 241 VLKIVEGTPSLPERYPKELNAIVESMLNKNPSLRPSAIEILKIPYLDLQNLMLCRYS 300  
DB 241 VLKIVEGTPSLPERYPKELNAIVESMLNKNPSLRPSAIEILKIPYLDLQNLMLCRYS 300  
QY 301 MTEDEKNDLCKEAAHIINAMQKRIHLOTLRALSEVQKMTPRERMLRKLQAADKARKL 360  
DB 301 MTEDEKNDLCKEAAHIINAMQKRIHLOTLRALSEVQKMTPRERMLRKLQAADKARKL 360  
QY 361 KKIIVEEKEENSKMQLSRNFQOLSVDVLEKTHLKGMEKEEQEGRGLSCSPQDEDE 420  
DB 361 KKIIVEEKEENSKMQLSRNFQOLSVDVLEKTHLKGMEKEEQEGRGLSCSPQDEDE 420  
QY 421 ERWQGREESDEPTLENLPESQIPSMDLHELESIVEDATSDLYH 466  
DB 421 ERWQGREESDEPTLENLPESQIPSMDLHELESIVEDATSDLYH 466

RESULT 5  
Q8C004 PRELIMINARY; PRT; 628 AA.  
ID Q8C004  
AC Q8C004;  
DT 01-MAR-2003 (TremBLrel. 23, Created)  
DT 01-MAR-2003 (TremBLrel. 23, Last sequence update)  
DT 01-OCT-2003 (TremBLrel. 25, Last annotation update)  
DE Similar to mRNA.  
DN 4932416N14RIK.  
OS Mus musculus (Mouse).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
OX NCBI\_TaxID=10090;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=C57BL/6J; TISSUE=Testis;  
RX MEDLINE=22354683; PubMed=12466851;  
RA The FANTOM Consortium.  
RA the RIKEN Genome Exploration Research Group Phase I & II Team;  
RT "Analysis of the mouse transcriptome based on functional annotation of

60,770 full-length cDNAs.";  
 RT NATURE 420:563-573 (2002).  
 RL EMBL; AK030042; BAC26756.1; -;  
 DR MGD; MGI:2442276; 4932416N14RIK.  
 DR GO; GO:0005524; F:ATP binding; IEA.  
 DR GO; GO:0004674; F:protein serine/threonine kinase activity; IEA.  
 DR GO; GO:0004713; F:protein-tyrosine kinase activity; IEA.  
 DR GO; GO:0006468; P:protein amino acid phosphorylation; IEA.  
 DR InterPro; IPR000719; Prot kinase.  
 DR InterPro; IPR008271; Ser\_thr\_pkin\_AS.  
 DR InterPro; IPR001245; Tyr\_pkinase.  
 DR Pfam; PF00069; pkinase; 1.  
 DR PRINTS; PR00109; TYRKINASE.  
 DR ProDom; PD000001; Prot kinase; 1.  
 DR SMART; SM00220; S\_TKc; 1.  
 DR SMART; SM00219; TyKc; 1.  
 DR PROSITE; PS00011; PROTEIN\_KINASE\_DOM; 1.  
 DR PROSITE; PS00108; PROTEIN\_KINASE\_ST; 1.  
 DR PROSITE; PS00108; PROTEIN\_KINASE\_ST; 1.  
 SQ SEQUENCE 628 AA; 71666 MW; AF6672BF80DCFE71 CRK64;  
 Query Match 70.3%; Score 1757.5; DB 11; Length 628;  
 Best Local Similarity 73.4%; Pred. No. 3.7e-110;  
 Matches 343; Conservative 37; Mismatches 50; Indels 37; Gaps 2;  
 QY 1 MLKFOBAKCVS-GSTAISTYKTLIARYVLOQKLGSGSGFTVYLVSDKKAKRGEELKV 59  
 DB 1 MLKFOETAKCVGRPTVIMPTALIAIRYVLOQKLGSGSGFTVYLVSDKKAKRGEELKV 60  
 QY 60 LKEISVGLNPNETVQANLEAQLSLDHPAIVKPHASFEVDNFCIIITEYCEGRDLDDK 119  
 DB 61 LKEISVGLNPNETVQANVEAQLSLRHPAIVRPHASFEQETFCIIITEYCEGRDLDDYR 120  
 QY 120 IQEYKQAGKIPFENQIIEWFIQLLGVDMYHERRILHDLKSNVFLKNLLKIGDFGV 179  
 DB 121 IQEYKAGKVFENQIIEWFIQLLGVDMYHERRILHDLKSNVFLKNLLKIGDFGV 180  
 QY 180 RLKMGSCDLATLTGTTPHYMSPEALKHQGYDTPKSDIWSLACILYEMCCMHAFAGSNFLS 239  
 DB 181 RLKMGSCDLATLTGTTPHYMSPEALKHQGYDTPKSDIWSLACILYEMCCMHAFAGSNFLS 240  
 QY 240 IVLKIVEGDTPSLPERYPKELNAIMESMLNKNPSLRPSAAILKAPYMEEQQLLMCKYP 299  
 DB 241 VLVNIVEGKTPSLPDYPRPRELNTIMERLNKSPSLRPSAAILKAPYMEEQQLLMCKYP 300  
 QY 300 EMTLEDKNDCKEAAHIINAMOKRIHLQTLRALSEVQKTPRERMRLKQAADKARK 359  
 DB 301 EMTLEDKNSVCKEAAHTINAVQKLHLQTLQALSQTKTTPRERMLRKLQAADERARR 360  
 QY 360 LKKIVBEKYEENSKRMQELSRNFQOLSVDVLHETHLKGMEKEEPGRSLSCSPQDED 419  
 DB 361 LKKIAENYKENDKRMQALSRNVGSHVHLH----- 393  
 QY 420 EERWQGRREESDEPTLENTPEISQIPSMDLHELSEIVEDATSDLG 466  
 DB 394 -----ELDERTLESPEQSLPCLDLDELEPSLEDTIVDLGHY 431

RESULT 6  
 Q8BL56 PRELIMINARY; PRT; 438 AA.  
 ID Q8BL56  
 AC Q8BL56;  
 DT 01-MAR-2003 (TrEMBLrel. 23, Created)  
 DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)  
 DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)  
 DE Similar to mRNA (Fragment).  
 GN 4932416N14RIK.  
 OS Mus musculus (Mouse).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 OC NCBI\_TaxID=10090;  
 RN [1]  
 RP SEQUENCE FROM N.A.

RC STRAIN=C57BL/6J; TISSUE=Cortex;  
 RX MEDLINE=22354683; PubMed=12466851;  
 RA The FANTOM Consortium,  
 RT the RIKEN Genome Exploration Research Group Phase I & II Team;  
 RA "Analysis of the mouse transcriptome based on functional annotation of  
 60,770 full-length cDNAs.";  
 RT NATURE 420:563-573 (2002).  
 DR EMBL; AK030042; BAC26756.1; -;  
 DR MGD; MGI:2442276; 4932416N14RIK.  
 DR GO; GO:0005524; F:ATP binding; IEA.  
 DR GO; GO:0004674; F:protein serine/threonine kinase activity; IEA.  
 DR GO; GO:0004713; F:protein-tyrosine kinase activity; IEA.  
 DR GO; GO:0006468; P:protein amino acid phosphorylation; IEA.  
 DR InterPro; IPR000719; Prot kinase.  
 DR InterPro; IPR002290; Ser\_thr\_pkinase.  
 DR InterPro; IPR008271; Ser\_thr\_pkin\_AS.  
 DR InterPro; IPR001245; Tyr\_pkinase.  
 DR Pfam; PF00069; pkinase; 1.  
 DR PRINTS; PR00109; TYRKINASE.  
 DR ProDom; PD000001; Prot kinase; 1.  
 DR SMART; SM00220; S\_TKc; 1.  
 DR SMART; SM00219; TyKc; 1.  
 DR PROSITE; PS00011; PROTEIN\_KINASE\_DOM; 1.  
 DR PROSITE; PS00108; PROTEIN\_KINASE\_ST; 1.  
 DR PROSITE; PS00108; PROTEIN\_KINASE\_ST; 1.  
 SQ SEQUENCE 438 AA; 50285 MW; DE540DB054F26D78 CRK64;  
 Query Match 70.1%; Score 1752.5; DB 11; Length 438;  
 Best Local Similarity 73.2%; Pred. No. 5.2e-110;  
 Matches 342; Conservative 37; Mismatches 51; Indels 37; Gaps 2;  
 QY 1 MLKFOBAKCVS-GSTAISTYKTLIARYVLOQKLGSGSGFTVYLVSDKKAKRGEELKV 59  
 DB 1 MLKFOETAKCVGRPTVIMPTALIAIRYVLOQKLGSGSGFTVYLVSDKKAKRGEELKV 60  
 QY 60 LKEISVGLNPNETVQANLEAQLSLDHPAIVKPHASFEVDNFCIIITEYCEGRDLDDK 119  
 DB 61 LKEISVGLNPNETVQANVEAQLSLRHPAIVRPHASFEQETFCIIITEYCEGRDLDDYR 120  
 QY 120 IQEYKQAGKIPFENQIIEWFIQLLGVDMYHERRILHDLKSNVFLKNLLKIGDFGV 179  
 DB 121 IQEYKAGKVFENQIIEWFIQLLGVDMYHERRILHDLKSNVFLKNLLKIGDFGV 180  
 QY 180 RLKMGSCDLATLTGTTPHYMSPEALKHQGYDTPKSDIWSLACILYEMCCMHAFAGSNFLS 239  
 DB 181 RLKMGSCDLATLTGTTPHYMSPEALKHQGYDTPKSDIWSLACILYEMCCMHAFAGSNFLS 240  
 QY 240 IVLKIVEGDTPSLPERYPKELNAIMESMLNKNPSLRPSAAILKAPYMEEQQLLMCKYP 299  
 DB 241 VLVNIVEGKTPSLPDYPRPRELNTIMERLNKSPSLRPSAAILKAPYMEEQQLLMCKYP 300  
 QY 300 EMTLEDKNDCKEAAHIINAMOKRIHLQTLRALSEVQKTPRERMRLKQAADKARK 359  
 DB 301 EMTLEDKNSVCKEAAHTINAVQKLHLQTLQALSQTKTTPRERMLRKLQAADERARR 360  
 QY 360 LKKIVBEKYEENSKRMQELSRNFQOLSVDVLHETHLKGMEKEEPGRSLSCSPQDED 419  
 DB 361 LKKIAENYKENDKRMQALSRNVGSHVHLH----- 393  
 QY 420 EERWQGRREESDEPTLENTPEISQIPSMDLHELSEIVEDATSDLG 466  
 DB 394 -----ELDERTLESPEQSLPCLDLDELEPSLEDTIVDLGHY 431

RESULT 7  
 Q8BW62 PRELIMINARY; PRT; 365 AA.  
 ID Q8BW62  
 AC Q8BW62;  
 DT 01-MAR-2003 (TrEMBLrel. 23, Created)  
 DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)  
 DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)  
 DE Similar to mRNA (Fragment).  
 GN 4932416N14RIK.

```
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Oviduct;
RX MEDLINE=22354683; PubMed=12466851;
RA The PANTOM Consortium.
RA "Analysis of the mouse transcriptome based on functional annotation of
RT 60,770 full-length cDNAs.";
RL Nature 420:563-573 (2002).
DR EMBL; AK054237; BAC35699.1; -
DR MGD; MGI:2442276; 4932416N14rik.
DR GO; GO:0005524; F:ATP binding; IEA.
DR GO; GO:0004674; F:protein serine/threonine kinase activity; IEA.
DR GO; GO:0004713; F:protein-tyrosine kinase activity; IEA.
DR GO; GO:0006468; P:protein amino acid phosphorylation; IEA.
DR InterPro; IPR000719; Prot kinase.
DR InterPro; IPR002290; Ser Thr pkinase.
DR InterPro; IPR008271; Ser Thr pkin_AS.
DR InterPro; IPR001245; Tyr_pkinase.
DR Pfam; PF00069; pkinase; 1.
DR PROSITE; PS00109; TYRKINASE.
DR ProDom; PD000001; Prot Kinase; 1.
DR SMART; SM00220; S_TKc; 1.
DR PROSITE; PS00111; PROTEIN_KINASE_DOM; 1.
DR PROSITE; PS00108; PROTEIN_KINASE_ST; 1.
FT NON_TER 365
SQ SEQUENCE 365 AA; 41752 MW; 01961635AY59B7F2 CRC64;

Query Match 63.8%; Score 1596.5; DB 11; Length 365;
Best Local Similarity 83.8%; Pred. No. 1.3e-99;
Matches 305; Conservative 29; Mismatches 29; Indels 1; Gaps 1;

QY 1 MLKFOEAKCVS-GSTAISTYPTKTLIARRYVLOQKLGSGSGFTVYLVSDKKAKGPEELKV 59
DB 1 MLKFOETAKCVGRPTVTPMYPTALIARRYVLOQKLGSGSGFTVYLVSDKKAKGPEELKV 60
QY 60 LKEISVGLNPNETVQANLEAQLSKLDHPATVKFHASVEODNFCITTEYCEGRDLDDK 119
DB 61 LKEISVGLNPNETVQANVEAQLSLRHLHPALVRFHAFSFWGETTCITTEYCEGRDLDYR 120
QY 120 IQEYQOAGKIFPENQIIEWFIOQLLGVDYMHERRILHRDLKSNVFLKXNLLKIGDFGV 179
DB 121 IQEYKEAGVFAENQIVENFIOQLLGVDYMHERRILHRDLKSNVFLKXNLLKIGDFGV 180
QY 180 RLIMGSCDLATLTGTTPHYMSPEALKHQGYDTKSDIWSLACILYEMCCMHAFAGSNFLS 239
DB 181 RLIMGSCDLATLTGTTPHYMSPEALKHQGYDAKSDIWSLACILYEMCCJDAHAFAGSNFLS 240
QY 240 IVLKIVEGDTPSLPERYPKELNAIMESMLNKNPSLRPSAIEILKIPYLDEQLQNLNLCRYS 299
DB 241 VVLNIVEGKTPSLPDYRPRELNTIMERLNKSPSLRPSAAILKAPYMEQQLQNLNLCYP 300
QY 300 EMTLEDXNLDCKEAAHIINAMQRIHLOTLRALSEVQKMTPRERMLRKLQAADKARKK 359
DB 301 EMTLEDXNSVCKEAAHTINAVQKKLHLOTLQALSDTQKTPRERMLRKLQAADKARKK 360
QY 360 LKKI 363
DB 361 LKKI 364

RESULT 8
QH5F4 PRELIMINARY; PRT; 463 AA.
AC QH5F4
DT 01-MAR-2001 (TrEMBLrel. 16, Created)
DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
```

```
DE Hypothetical protein FLJ23495.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Lung;
RA Kawakami T., Noguchi S., Itoh T., Shigeta K., Senba T., Matsumura K.,
RA Nakajima Y., Mizuno T., Morinaga M., Tanigami A., Fujiwara T., Ono T.,
RA Yamada K., Fujii Y., Ozaki K., Hirao M., Ohmori Y., Ota T., Suzuki Y.,
RA Obayashi M., Nishi T., Shibahara T., Tanaka T., Nakamura Y.,
RA Isogai T., Sugano S.;
RT "NEDO human cDNA sequencing project.";
RL Submitted (AUG-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AK027148; BAB15672.1; -
DR GO; GO:0005524; F:ATP binding; IEA.
DR GO; GO:0004672; F:protein kinase activity; IEA.
DR GO; GO:0016740; F:transferase activity; IEA.
DR GO; GO:0006468; P:protein amino acid phosphorylation; IEA.
DR InterPro; IPR000719; Prot kinase.
DR Pfam; PF00069; pkinase; 1.
DR ProDom; PD000001; Prot kinase; 1.
DR PROSITE; PS00111; PROTEIN_KINASE_DOM; 1.
DR KW Hypothetical protein; ATP-binding; Transferase.
SQ SEQUENCE 463 AA; 53403 MW; DEB585A8C1C939F9 CRC64;

Query Match 59.2%; Score 1480; DB 4; Length 463;
Best Local Similarity 100.0%; Pred. No. 1.2e-91;
Matches 284; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 183 MGSCDLATLTGTTPHYMSPEALKHQGYDTKSDIWSLACILYEMCCMHAFAGSNFLSIVL 242
DB 1 MGSCDLATLTGTTPHYMSPEALKHQGYDTKSDIWSLACILYEMCCMHAFAGSNFLSIVL 60
QY 243 KIVEGDTPSLPERYPKELNAIMESMLNKNPSLPSAIEILKIPYLDEQLQNLNLCRYSEM 302
DB 61 KIVEGDTPSLPERYPKELNAIMESMLNKNPSLPSAIEILKIPYLDEQLQNLNLCRYSEM 120
QY 303 LEDKNLDCQKEAAHIINAMQRIHLOTLRALSEVQKMTPRERMLRKLQAADKARKLKK 362
DB 121 LEDKNLDCQKEAAHIINAMQRIHLOTLRALSEVQKMTPRERMLRKLQAADKARKLKK 180
QY 363 IVEKYEENSKRMQELSRNFQQLSVDLVHEKHLKGMEKEQPEGRSLCSPODEDEER 422
DB 181 IVEKYEENSKRMQELSRNFQQLSVDLVHEKHLKGMEKEQPEGRSLCSPODEDEER 240
QY 423 WQGREESDEPTLENLPESQIPISMDLHELESIVEDATSDLYH 466
DB 241 WQGREESDEPTLENLPESQIPISMDLHELESIVEDATSDLYH 284

RESULT 9
QH5NT6 PRELIMINARY; PRT; 395 AA.
ID QH5NT6
AC QH5NT6
DT 01-MAR-2002 (TrEMBLrel. 20, Created)
DT 01-MAR-2002 (TrEMBLrel. 20, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE mRNA, similar to human hypothetical protein FLJ23495, complete cds.
OS Macaca fascicularis (Crab eating macaque) (Cynomolgus monkey).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Cercopitheidae;
OC Cercopitheidae; Macaca.
OX NCBI_TaxID=9541;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Testis;
RA Hashimoto K., Oasda N., Hida M., Kusuda J., Tanuma R., Hirai M.,
RA Terao K., Suzuki Y., Sugano S.;
RT "Isolation of novel full-length cDNA clones from macaque testis cDNA
libraries.";
```



RL Submitted (JUL-2001) to the EMBL/GenBank/DBJ databases.  
 DR EMBL: AB066559; BAB84034.1; -.  
 DR GO: GO:0005524; F-ATP binding; IEA.  
 DR GO: GO:0004672; F-protein kinase activity; IEA.  
 DR GO: GO:0016740; F-transferase activity; IEA.  
 DR GO: GO:0006468; P-protein amino acid phosphorylation; IEA.  
 DR InterPro: IPR000719; Prot kinase.  
 DR Pfam: PF00069; pkinase; 1.  
 DR ProDom: PD000001; Prot kinase; 1.  
 DR PROSITE: PS00011; PROTEIN\_KINASE\_DOM; 1.  
 KW ATP-binding; Transferase.  
 SQ SEQUENCE 395 AA; 45391 MW; 585EF13B2B89B85 CRC64;  
 Query Match 55.8%; Score 1396; DB 6; Length 395;  
 Best Local Similarity 94.7%; Pred. No. 4.6e-86;  
 Matches 269; Conservative 9; Mismatches 9; Indels 0; Gaps 0;  
 YQ 183 MGSCDLATLTGPHYMSPEALKHQGYDYSKDIWSLACILYEMCCNHFAGSNFLSIVL 242  
 Db 1 MGSCDLATLTGPHYMSPEALKHQGYDYSKDIWSLACILYEMCCNHFAGSNFLSIVL 60  
 YQ 243 KIVEGTPSLPERYPKELNAIMESMLNKNPSLRPSAIELKIPYLDLQNLNLCRYSEMT 302  
 Db 61 KIVEGTPSLPERYPKELNAIMESMLNKNPSLRPSAIELKIPYLDLQNLNLCRYSEMT 120  
 YQ 303 LEDKNDCKEAAHINAMQRIHLQTLRALSEVQQTPTPRMRLKQAADKARKLK 362  
 Db 121 LEDKNDCKEAAHINAMQRIHLQTLRALSEVQQTPTPRMRLKQAADKARKLK 180  
 YQ 363 IVEKEEENSKMOELRSNFQOLSVDVLEHKLGMKEEKEOPEGRSCSPQDEDEER 422  
 Db 181 IVEKEEENSKMOELRSNFQOLSVDVLEHKLGMKEEKEOPEGRSCSPQDEDEER 240  
 YQ 423 WQGREESDEPTLENIPESQIPSPMDLHELESIVEDATSDLYGH 466  
 Db 241 WQGREESDEPTLENIPESQIPSPMDLHELESIVEDATSDLYGH 284  
 RESULT 10  
 Q7T299 PRELIMINARY; PRT; 849 AA.  
 AC Q7T299;  
 DT 01-OCT-2003 (T-EMBLrel. 25, Created)  
 DT 01-OCT-2003 (T-EMBLrel. 25, Last sequence update)  
 DT 01-OCT-2003 (T-EMBLrel. 25, Last annotation update)  
 DE Hypothetical protein.  
 OS Brachydanio rerio (Zebrafish) (Danio rerio).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes;  
 OC Cyprinidae; Danio.  
 OX NCBI\_TaxID=7955;  
 [1]  
 RN SEQUENCE FROM N.A.  
 RC TISSUE=Kidney;  
 RX MEDLINE=22388257; PubMed=12477932;  
 RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,  
 RA Klausner R.D., Collins P.S., Wagner L., Shenmen C.M., Schuler G.D.,  
 RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,  
 RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,  
 RA Diatchenko L., Marusina K., Farmer A., Rubin G.M., Hong L.,  
 RA Stapleton M., Soares M.B., Donald M.F., Casavant T.L., Scheetz T.E.,  
 RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,  
 RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,  
 RA Bosak S.A., McSwain P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,  
 RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,  
 RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,  
 RA Fahey J., Heiton E., Kettelman M., Madan A., Rodrigues S., Sanchez A.,  
 RA Whitting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,  
 RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,  
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,  
 RA Krzywinski M.I., Skalska U., Smailus D.E., Schnerch A., Schein J.E.,  
 RA Jones S.J., Marra M.A.;  
 "Generation and initial analysis of more than 15,000 full-length human  
 and mouse cDNA sequences.";  
 Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).  
 [2]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Kidney;  
 RA Strausberg R.;  
 RL Submitted (JUL-2003) to the EMBL/GenBank/DBJ databases.  
 DR EMBL: BC054633; AAH54633.1; -.  
 KW Hypothetical protein.  
 SQ SEQUENCE 849 AA; 95966 MW; E4EBB117A0147D7B CRC64;  
 Query Match 23.7%; Score 593.5; DB 13; Length 849;  
 Best Local Similarity 29.3%; Pred. No. 1.3e-11;  
 Matches 141; Conservative 92; Mismatches 178; Indels 71; Gaps 11;  
 YQ 29 YVLOQKLGSGFVTYLVSDKKAKRGBELKVLKEISVGLNPNETVQANLEAQLSKLDH 88  
 Db 4 YLFRVVGKSGYEVNLRHKSDBK--QVVIKLNLRITSSRRERRAAQEAQLLSQLKH 60  
 YQ 89 PAIVKHFASVFOO-NFCIITEYCEGRDLDKIOEYKQAKIPENQIIEWFIQLLQVD 147  
 Db 61 FNIWVRESWEGEDCOLYIVMGFCGGDLVHRLKQ--QKGEILPERQVVEWFIQIAMAQ 118  
 YQ 148 YMHERRTILHRDLKSKNVL-KNNLLKIGDFGVSRLMGSCDLATLTGPHYMSPEALKH 206  
 Db 119 YLHEKHILHRDLKTONIFLTKTNIIVKGDGLIARVLENQNDMASTLIGTFYMSPLFSN 178  
 YQ 207 QGYDTKSDIWSLACILYEMCCNHFAGSNFLSIVLKIYEGDTPSLPERYPKELNAIMES 266  
 Db 179 KPYNKSDVMALGCCVYEMATLKAFNAKDMNLAVRIVEGKLPMPSKYDPOLGELIKR 238  
 YQ 267 MLNKNPSLRPSAIELKIPYLDLQNLNLCRYSEMT----- 302  
 Db 239 MLCKKPEDRPDVKHIDRQPIKHQISNLFLEATKTAKRKNNAOKLSAGSDATKKN 298  
 YQ 303 -----LEDKNDCKEAAHINAMQ-----RIHLQTLRALSEVQQTPTPR 344  
 Db 299 QVWQPOCLNSESCTCGKKAEEIYLNQKPCNGAWENAVPKHMPKSPTRDIENSTGSI 358  
 YQ 345 MRLKLOAADEKARKLKIVVEEYKVN-----SKRQELSRNFQOLSVDVLEHKLHG 399  
 Db 359 ATISNIDIEIQKQKAKPPKPPSHQNNLPSVSKR-----REKDPGAPQTHPHKQVSG 412  
 YQ 400 MEEKEOPEGRKL-SCSPQDEDEERWQGREES-----EPTLENLPES-----QIPSPMD 448  
 Db 413 VGGTDEKMSANASSIIPKPADRTKMPNKSALDVSIDIKDDTKMLQEAQVQDLPVEPTE 472  
 YQ 449 LH 450  
 Db 473 LH 474  
 RESULT 11  
 Q8CD72 PRELIMINARY; PRT; 424 AA.  
 AC Q8CD72;  
 DT 01-MAR-2003 (T-EMBLrel. 23, Created)  
 DT 01-MAR-2003 (T-EMBLrel. 23, Last sequence update)  
 DT 01-OCT-2003 (T-EMBLrel. 25, Last annotation update)  
 DE NINA.  
 GN NEKI.  
 OS Mus musculus (Mouse).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 OX NCBI\_TaxID=10090;  
 [1]  
 RN SEQUENCE FROM N.A.  
 RC STRAIN=CS7BL/6J; TISSUE=Testis;  
 RX MEDLINE=22354683; PubMed=12466851;  
 RA The FANTOM Consortium,  
 RA the RIKEN Genome Exploration Research Group Phase I & II Team;  
 RT "Analysis of the mouse transcriptome based on functional annotation of  
 60,770 full-length cDNAs.";

```
RL Nature 420:563-573 (2002).
DR EMBL; AK031330; BAC27350.1; -.
DR MGI; MGI:97303; Nekl.
DR GO; GO:0005524; F:ATP binding; IEA.
DR GO; GO:0004674; F:protein serine/threonine kinase activity; IEA.
DR GO; GO:0004713; F:protein-tyrosine kinase activity; IEA.
DR GO; GO:0004688; P:protein amino acid phosphorylation; IEA.
DR InterPro; IPR000719; Prot_kinase.
DR InterPro; IPR002290; Ser_thr_pkinase.
DR InterPro; IPR008271; Ser_thr_pkin_AS.
DR Pfam; PF00069; pkinase; 1.
DR PRINTS; PR00109; TYRKINASE.
DR ProDom; PD000001; Prot_kinase; 1.
DR SMART; SM00220; S_TK; 1.
DR SMART; SM00219; TYRK; 1.
DR PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.
DR PROSITE; PS00111; PROTEIN_KINASE_DOM; 1.
DR PROSITE; PS00108; PROTEIN_KINASE_ST; 1.
SQ SEQUENCE 424 AA; 48635 MW; 416AD0E55C13CEFE CRC64;

Query Match 23.6%; Score 589.5; DB 11; Length 424;
Best Local Similarity 30.3%; Pred. No. 1e-31;
Matches 132; Conservative 88; Mismatches 128; Indels 87; Gaps 10;

QY 28 RYVLOKLGSGFQTVLYSDKAKRGELKVLKEISVGEINPNETVQANLEAQLLSKLD 87
DB |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 3 KYVRLQKIGEGFGKAVLV--KSTEDGRHY-VIKESINSRMSDKROSRREVAVLANK 59
DB |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 88 HPAIVKHFASFVEQDNFCIITYCEGRDLDDKIQYKQAGKIFPENQIIEFIQLLQVD 147
DB |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 60 HENIVQYKESFENGSLYIVMDYCEGGDLFKEN--AQKALFQEDQLDWFVQICLALK 117
DB |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 148 YMHERILHRDLKSNVFL-KNNLLKIGDGVSRILMGSCDLATLTGTTPHYMSPEALKH 206
DB |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 118 HVHDKILHRDIKSNQIFLTKDGTQGLDGFGLRVNLSTVELAR-CIGTPYVLSPEICEN 177
DB |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 207 QGVYDKSDIWSLACILYECMNHAFAGSNFSLVILKIVEGDTPSLPERYPKELNAIMES 266
DB |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 178 KPYNKSIDLWLCGVLYELCTKHAFAGMKNVLKLIISGFPVPVSHYDLRLSLSQ 237
DB |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 267 MLNKNPSLRPSAIEILKIPYLDQQLNLM-----C----- 296
DB |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 238 LFKRNPRDRPSNVSILEKGFIAKRIEKLPSLPOLIAEFCLKTSKFGPQLPGKRPASGQ 297
DB |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 297 -----RYSEMTLEDKILDCQKEAAHILINAMQKRIHLQIR 331
DB |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 298 GYSSVFPAQKITKPAKYGVPLTYKYGDKLLEKPPPKHQAHOI----- 344
DB |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 332 ALSEVOKMTPRMRRLKLAADKARKLKKTVEEYKSNRMQELSRNFPQQLSVDVL 391
DB |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 345 ---PVKKNSSGERKQWSEAA--KKRLFEIEKEK-----KQKQIRFLKAEQMK--R 391
DB |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 392 HEKTHLKGMEKEEQ 406
DB |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 392 QSKQRLERINRAREQ 406
DB |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:

RESULT 12
Q8CCJ0 PRELIMINARY; PRT; 614 AA.
AC Q8CCJ0;
DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Weakly similar to protein kinase nekl.
GN BB049667.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.

RC STRAIN=C57BL/6J; TISSUE=Cerebellum;
RX MEDLINE=22354683; PubMed=12466851;
RA The FANTOM Consortium,
RA "The RIKEN Genome Exploration Research Group Phase I & II Team;
RT "Analysis of the mouse transcriptome based on functional annotation of
RL Nature 420:563-573 (2002).";
DR EMBL; AK032672; BAC27980.1; -.
DR MGI; MGI:2142824; BB049667.
DR GO; GO:0005524; F:ATP binding; IEA.
DR GO; GO:0004674; F:protein serine/threonine kinase activity; IEA.
DR GO; GO:0004713; F:protein-tyrosine kinase activity; IEA.
DR GO; GO:0004688; P:protein amino acid phosphorylation; IEA.
DR InterPro; IPR000719; Prot_kinase.
DR InterPro; IPR002290; Ser_thr_pkinase.
DR InterPro; IPR008271; Ser_thr_pkin_AS.
DR Pfam; PF00069; pkinase; 1.
DR PRINTS; PR00109; TYRKINASE.
DR ProDom; PD000001; Prot_kinase; 1.
DR SMART; SM00220; S_TK; 1.
DR SMART; SM00219; TYRK; 1.
DR PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.
DR PROSITE; PS00111; PROTEIN_KINASE_DOM; 1.
DR PROSITE; PS00108; PROTEIN_KINASE_ST; 1.
SQ SEQUENCE 614 AA; 70032 MW; DA95120EA45D71BC CRC64;

Query Match 23.5%; Score 586.5; DB 11; Length 614;
Best Local Similarity 31.8%; Pred. No. 2.6e-31;
Matches 148; Conservative 91; Mismatches 145; Indels 81; Gaps 16;

QY 35 LGSSGFTGVLYSDKAKRGELKVLKEISV---GELNPNETVQANLEAQLLSKLDHPAI 91
DB |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 10 IGESTFRKVLAKDXE---SSHCVIKEISLTKKEASKNEVI-----LLAEMEHPNI 59
DB |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 92 VKHFASVFEQDNFCIITYCEGRDLDDKIQYKQAGKIFPENQIIEFIQLLQVDVYHME 151
DB |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 60 VTFPSSFOENGRLFIVMYECDDGLMORIQ--RQRGVWFSEDLQLCWFVQISLGLKHIHD 117
DB |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 152 RRIILHRDLKSNVFLKNN--LLKIGDGVSRILMGSCDLATLTGTTPHYMSPEALKHQGY 209
DB |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 118 RKILHRDIKSNQIFLTKGKNGVAKLGDFTARTLNDSELAQTCACTAGTTPYLSPEICQNRPY 177
DB |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 210 DTKSDIWSLACILYECMNHAFAGSNFSLVILKIVEGDTPSLPERYPKELNAIMESMLN 269
DB |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 178 NNTKDIWSLGVLYELCTKHAFAGMKNVLKLIISGFPVPVSHYDLRLSLQ 237
DB |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 270 KNPRLPSAIEILKIPYLDQQLNLMCR--YSEMTLEDKILDCQKEA--AHINA----- 320
DB |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 238 VSPQDRFESVTSLLKRPF---LETLIARSLYPEV-----CSRRTQSHAHMENMAIGP 285
DB |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 321 -----MOKRIHLQTLRALSEVQ-----KMTPRRMELR-----KLQAA 353
DB |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 286 TACWVSPFWSNAYLORKFEAQYKLVKVERQLGLRPPSVVEHPNEGKQLQSHWEETKFQEL 345
DB |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 354 DEKARKLK-----KIVEEYKSNRMQELSRNFPQQLSVDVLHEKTHLKGMEF-----KE 404
DB |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 346 QYRKNKMDQBYWKQLEIRIQQVYHNDMKEIKKMGRELKRVKVFELSLDKISEEDTVQE 405
DB |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 405 BQPSGLSCSPQDED-----EERWQGREESDEPTLENL--PESQ 442
DB |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 406 NEAVDKLNATLSFEDGTGKFOHRCHEEHEDYTDRAFEELCGPEAE 450
DB |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:

RESULT 13
Q7TSC3 PRELIMINARY; PRT; 627 AA.
AC Q7TSC3;
DT 01-OCT-2003 (TrEMBLrel. 25, Created)
DT 01-OCT-2003 (TrEMBLrel. 25, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Hypothetical protein.
OS Mus musculus (Mouse).
```

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 OC NCBI\_TaxID=10090;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Olfactory epithelium;  
 RX MEDLINE=2238257; PubMed=12477932;  
 RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G., Schuler G.D.,  
 RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,  
 RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,  
 RA Hopkins R.F., Jordan L., Moore T., Max S.I., Wang J., Hong F.,  
 RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hsieh L.,  
 RA Stapleton M., Soares M.B., Bonaldo M.P., Casavant T.L., Scheetz T.E.,  
 RA Brownstein M.J., Ueding T.B., Toshiyuki S., Abramson R.D., Mullany S.J.,  
 RA Rana S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullany S.J.,  
 RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,  
 RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,  
 RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,  
 RA Fahey J., Helton E., Kettaman M., Madan A., Rodriguez S., Sanchez A.,  
 RA Whitting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,  
 RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,  
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,  
 RA Krzywinski M.I., Skalska U., Smailus D.E., Schnerch A., Schein J.E.,  
 RA Jones S.J., Marra M.A.;  
 RT "Generation and initial analysis of more than 15,000 full-length human  
 and mouse cDNA sequences";  
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).  
 [2]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Olfactory epithelium;  
 RA Strausberg R.;  
 RL Submitted (JUN-2003) to the EMBL/GenBank/DBJ databases.  
 DR EMBL: BC053516; AAH53516.1; --  
 KW Hypothetical protein.  
 SQ SEQUENCE 627 AA; 17190 MW; AD5789B6621AC2P CRC64;  
 Query Match 23.5%; Score 586.5; DB 11; Length 627;  
 Best Local Similarity 31.8%; Pred. No. 2.7e-31;  
 Matches 148; Conservative 91; Mismatches 145; Indels 81; Gaps 16;  
 QY 35 LGSGSGFTVYLVSDKKAKGEELKVLKEISV---GELNPNETVQANLEAQLSLKLDHAI 91  
 DB 10 IGBGTGKYLAKDKSE---SHCVIKELSLKKEAKSEVI-----LLARMEFNI 59  
 QY 92 VKPHASFEVDNFCIIITEYCEGRDLDDKIQYKQAGKIPFENQIIEWFIQLLGVYDYMHE 151  
 DB 60 VTFPFSFQENGRFLIVMEYCDGDLQRIQ---RQRGVMPSEQILCWFTVLSGLKHIHD 117  
 QY 152 RRLHRLDKSKNVFLKVN---LLKIGDGVSRLLMGSCLDATTLTGTPHYMSPEALKHQGY 209  
 DB 118 RKILHRDIKSONIFLSKNGWAKLGDGFGPARTLNDMSLAEQAOTCACTPYLSEICQNPY 177  
 QY 210 DTKSDIWSLACILYEMCCMNHAFAGSNFLSIVLKIIVEGDTPLSPRYPKELNAINMESLN 269  
 DB 178 NNKTDIWSLGCVLVELCTLKHFPESNFHVLKICQGEVAPISPHFSDLSQSLPQLFR 237  
 QY 270 KNPSPRPSAIEILKIPYLDQNLNLCR---YSEMTLEDKNLDCQKEA---AHINA----- 320  
 DB 238 VSPQDRPSVTSLLKRPF---LETIARSLYPEV-----CSRRIOSHAHMENNAIGP 285  
 QY 321 -----MOKRIHLQTLRALSEVO-----KMTPRERMRLR-----KLQAA 353  
 DB 286 TACWRVSPSAAYLQKFAQYKLVKVERQLGLRSPSSVEPHNEGEKQLQSHWEETKFQEL 345  
 QY 354 DEKARKLK-----KIVVEKYENSKRMQELSRNFQQLSDVVLHKEKTLKGNWE-----KE 404  
 DB 346 QYRKNNKMDQYWKQLEIRQYVHNDMEIKKMGRELKRVVVKFPISLDKCITSEEDTVQE 405  
 QY 405 EQPEGSLSCSPQDED-----ERRWQGRREESDEPTLENI---PESO 442  
 DB 406 NEADVKNATLSFEDGTKEQHRCKEHEHEDYTDRAFEELCGPEAE 450

RESULT 14  
 Q8J023  
 ID Q8J023 PRELIMINARY; PRT; 489 AA.  
 AC Q8J023;  
 DT 01-MAR-2003 (TREMBlrel. 23, Created)  
 DT 01-MAR-2003 (TREMBlrel. 23, Last sequence update)  
 DT 01-OCT-2003 (TREMBlrel. 25, Last annotation update)  
 DE NIMA-related protein kinase 3.  
 GN NEK3.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.  
 NCBI\_TaxID=9606;  
 [1]  
 RP SEQUENCE FROM N.A.  
 RA Kimura M., Okano Y.;  
 RL "Molecular cloning of human NEK3";  
 RL Submitted (OCT-2001) to the EMBL/GenBank/DBJ databases.  
 DR EMBL: AB072828; BAC15599.1; --  
 DR GO: GO:0005524; F:ATP binding; IEA.  
 DR GO: GO:0004874; F:protein serine/threonine kinase activity; IEA.  
 DR GO: GO:0004713; F:protein-tyrosine kinase activity; IEA.  
 DR GO: GO:0006468; P:protein amino acid phosphorylation; IEA.  
 DR InterPro: IPR000719; Prot kinase.  
 DR InterPro: IPR002290; Ser/thr\_pkinase.  
 DR InterPro: IPR002271; Ser/thr\_pkinase.  
 DR InterPro: IPR001245; Tyr\_pkinase.  
 DR Pfam: PF00069; pkinase;  
 DR PRINTS: PRO0109; TYRKINASE.  
 DR PRODOM: PD000001; Prot\_kinase; 1.  
 DR SMART: SM00220; S\_TKc; 1.  
 DR SMART: SM00219; TyKc; 1.  
 DR PROSITE: PS00107; PROTEIN\_KINASE\_ATP; 1.  
 DR PROSITE: PS00011; PROTEIN\_KINASE\_DOM; 1.  
 DR PROSITE: PS00108; PROTEIN\_KINASE\_ST; 1.  
 KW Kinase.  
 SQ SEQUENCE 489 AA; 55983 MW; FAESDAAC01EFA88A CRC64;  
 Query Match 23.4%; Score 585.5; DB 4; Length 489;  
 Best Local Similarity 31.4%; Pred. No. 2.3e-31;  
 Matches 151; Conservative 87; Mismatches 154; Indels 89; Gaps 15;  
 QY 29 YVLOQKLGSGFTVYLVSDKKAKGEELKVLKEISVGEINPNETVQANLEAQLSLKLDH 88  
 DB 4 YVLMRMIGEGFGRALLVQHSS---NQMPAMKEILPKFSFN-TQNSKEAVLLAKMKH 59  
 QY 89 PAIVKPHASFEVDNFCIIITEYCEGRDLDDKIQYKQAGKIPFENQIIEWFIQLLGVYD 148  
 DB 60 PNIVAFKSEFEAGHLIVMEYCDGDLQRIQ---QKGLFPEDMLNWFQWCLGVNH 117  
 QY 149 MHERRILHRDLKSKNVFL-KNNLLKIGDGVSRLLMGSCLDATTLTGTPHYMSPEALKHQ 207  
 DB 118 IHKRVLRHDIKSKNIFLTQNGKVLGDFGSARLLSNPMAFACVYVGTPTVYVPEIWEHL 177  
 QY 208 GYDTKSDIWSLACILYEMCCMNHAFAGSNFLSIVLKIIVEGDTPLSPRYPKELNAINMES 267  
 DB 178 PYNKSDIWSLGCVLVELCTLKHFPFQANSWNKLVKVCQCISPLFSHYSELOFLVKQ 237  
 QY 268 LNKNPSPRPSAIEILKIPYLDQNLNLCR---YSEMTLEDKNLDCQKEA---AHINA-----DC 310  
 DB 238 FKRNPSHRPSATLLSGRIVARLVQKLPPEIIMEYGEVLEIKNSKNTPRKQDEEQ 297  
 QY 311 OKEAHI-INAMOKRIHLQTLRALSEVOQNTPEYVRLKLAQADEKARKLKVIEE--- 366  
 DB 298 DRKGSHTDLESINENLVESALRRVNEEK---GNKSVHLKRASSPNLHRKQWKNQVNTAL 355  
 QY 367 -----TALNASILTSLSAEDDRGSGVIKYSKNTTRKQWLKETPDTLLNLKNADLSLAFQTV 415  
 DB 356 TALENATLSLSAEDDRGSGVIKYSKNTTRKQWLKETPDTLLNLKNADLSLAFQTV 415  
 QY 388 VDLVHEKTLKG---MEEKES-----OPEGRSLSCSPQDEDEERWQGRREESDE 432  
 DB 416 IYRPGSEGFLKGPLSEBETASDSVDCGHSDVILDP-ELEPGDDEEDTD-----FEEEDN 470

Search completed: October 1, 2004, 11:20:45  
Job time : 125 secs

```
QY 433 P 433
Db 471 P 471

RESULT 15
Q8BSB6 PRELIMINARY; PRT; 302 AA.
AC Q8BSB6
DT 01-WAR-2003 (TREMELrel. 23, Created)
DT 01-WAR-2003 (TREMELrel. 23, Last sequence update)
DT 01-OCT-2003 (TREMELrel. 25, Last annotation update)
DE NIMA.
GN NEKI.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
SEQUENCE FROM N.A.
RP STRAIN=C57BL/6J; TISSUE=Embryo;
RX MEDLINE=22354693; PubMed=12466851;
RA The FANTOM Consortium.
RA the RIKEN Genome Exploration Research Group Phase I & II Team;
RT "Analysis of the mouse transcriptome based on functional annotation of
RL Nature 420:563-573(2002)."
DR EMBL; AK034754; BAC28822.1; -.
DR MGD; MGI:97303; Neki.
DR GO; GO:0005524; P:ATP binding; IEA.
DR GO; GO:0004674; P:protein serine/threonine kinase activity; IEA.
DR GO; GO:0004713; P:protein-tyrosine kinase activity; IEA.
DR GO; GO:0004688; P:protein amino acid phosphorylation; IEA.
DR InterPro; IPR000719; Prot_kinase.
DR InterPro; IPR002290; Ser_thr_kinase.
DR InterPro; IPR008271; Ser_thr_kinase.
DR InterPro; IPR001245; Tyr_kinase.
DR Pfam; PF00069; pkinase; 1.
DR PRINTS; PR00109; TYRKINASE.
DR ProDom; PD000001; Prot_kinase; 1.
DR SMART; SM00220; S_TKc; 1.
DR SMART; SM00219; TyKc; 1.
DR PROSITE; PS00107; PROTEIN KINASE ATP; 1.
DR PROSITE; PS00111; PROTEIN KINASE DOM; 1.
DR PROSITE; PS00108; PROTEIN KINASE_ST; 1.
SQ SEQUENCE 302 AA; 34328 MW; 6D5B14D7809CB341 CRC64;

Query Match 23.1%; Score 577; DB 11; Length 302;
Best Local Similarity 38.5%; Pred.No. 4.7e-31;
Matches 110; Conservative 71; Mismatches 95; Indels 10; Gaps 5;

QY 28 RYVLOKLGSGGTGYLYSDKKRGEELKVLKETSVCGLNPNETVQANLEAQLLSKLD 87
Db 3 KYVRLQIGEGSGKAVLV--KSTEDGRHY-VIKEINISRMDSKQESRRVAVLANWK 59
QY 88 HPAIVKPHASFEVDNFCITTEYCEGRDLDDKIQEYKQAGKIPPENQIIWFQILLGVD 147
Db 60 HPNIVQKESFEENGSLYIVMDYCEGGDLFXRN--AQKALFQEDQILDWFVQICLALK 117
QY 148 YMERILHRLDKSKNVFL-KNNLLKIGDGVSRLLMGSCDLATLTGTGPHYMSPEALKH 206
Db 118 HVHDKILHRLDKSKNVFLKDGTVQFGDFTARVLNSTVELARTCTGTPYLSPEICEN 177
QY 207 QGYDTKDSIWSLACILYEMCCMHAFAGNSFLSIVLKIVEGDTPLSPERYPKELNAIMES 266
Db 178 KPTNNKSDIHALGCVLYELCTLKHAFAGNMKNLVLIISGSPFPVSPHYSDLRSLLSQ 237
QY 267 MLNKNPSLRPSAIEILKIPYLDQLQNL-----CRYSEMTLEDKNL 308
Db 238 LFKRNPDRPSVNSILEKFIKRIKFLSPQVCVTLPLTLNCKVL 283
```

**This Page Blank (uspto)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: ~ October 2, 2004, 11:43:43 ; Search time 5789 Seconds  
(without alignments)  
10848.861 Million cell updates/sec

Title: US-10-803-278-3

Perfect score: 1449

Sequence: 1 atgtcgaaatcccaagagc.....actggaaaaatgaaaaataa 1449

Scoring table: IDENTITY\_NUC

Gapop 10.0 , Gapext 1.0

Searched: 3470272 seqs, 21671516995 residues

Total number of hits satisfying chosen parameters: 6940544

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

GenEmbl :

1: gb\_ba.\*

2: gb\_htg.\*

3: gb\_in.\*

4: gb\_om.\*

5: gb\_ov.\*

6: gb\_pat.\*

7: gb\_ph.\*

8: gb\_pl.\*

9: gb\_pr.\*

10: gb\_ro.\*

11: gb\_scs.\*

12: gb\_sy.\*

13: gb\_un.\*

14: gb\_vl.\*

15: em\_ba.\*

16: em\_fun.\*

17: em\_hum.\*

18: em\_in.\*

19: em\_mu.\*

20: em\_om.\*

21: em\_or.\*

22: em\_ov.\*

23: em\_pat.\*

24: em\_ph.\*

25: em\_pi.\*

26: em\_ro.\*

27: em\_sts.\*

28: em\_un.\*

29: em\_vl.\*

30: em\_htg\_hum.\*

31: em\_htg\_inv.\*

32: em\_htg\_other.\*

33: em\_htg\_mus.\*

34: em\_htg\_pln.\*

35: em\_htg\_rod.\*

36: em\_htg\_mam.\*

37: em\_htg\_vrt.\*

38: em\_sy.\*

39: em\_htgo\_hum.\*

40: em\_htgo\_mus.\*

41: em\_htgo\_other.\*

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	% Match	Query Length	DB ID	Description
1	1449	100.0	1854	9	BC028587 Homo sapi
2	1399.8	96.6	2080	9	AB071997 Homo sapi
3	1399	96.5	1938	6	AX166555 Sequence
4	1399	96.5	2552	6	AX746174 Sequence
5	1399	96.5	2768	9	HS0604785
6	1399	96.5	2939	9	AB071996 Homo sapi
7	1335	92.1	2936	9	AB064997 Macaca fa
8	918	63.4	2049	9	AK027148 Homo sapi
9	912.2	63.0	2902	9	AB066559 Macaca fa
10	651.8	45.0	2189	9	AK123021 Homo sapi
C 11	172.2	11.9	154330	2	AC025148 Homo sapi
C 12	172.2	11.9	133425	9	AC055733 Homo sapi
C 13	167.2	11.5	60398	9	AC116424 Homo sapi
C 14	153.4	10.6	111568	9	AC083908 Homo sapi
C 15	153.4	10.6	175923	2	AC064839 Homo sapi
C 16	152.2	10.5	1547	3	AK114791 Ciona int
17	148	10.2	2603	9	BC063044 Homo sapi
18	148	10.2	3698	9	HUMSTK2A
19	147.8	10.2	1510	6	AX657954 Sequence
20	147.8	10.2	1578	6	AX657957 Sequence
21	147.8	10.2	1581	6	AX657962 Sequence
22	147.8	10.2	1610	9	AB072828 Homo sapi
23	147.8	10.2	1638	6	AX803387 Sequence
24	147.8	10.2	2059	6	BD251812 Phosphory
25	147.8	10.2	2110	6	AX504240 Sequence
26	147.8	10.2	2134	9	BC019916 Homo sapi
27	146.4	10.1	2463	10	MMSTK2S
28	146.4	10.1	2607	10	MMSTK2L
29	146.4	10.1	2715	6	AX657960 Sequence
30	146.4	10.1	2715	6	AX659642 Sequence
31	146.4	10.1	2715	10	AF099067 Mus muscu
32	146.4	10.1	2778	10	BC057939 Mus muscu
33	146.2	10.1	1868	9	HSNEK3R
34	146.2	10.1	2364	6	AX405633 Sequence
35	146.2	10.1	4263	6	AR223879 Sequence
36	146.2	10.1	4263	10	S45828 nek1=serine
37	143.2	9.9	3645	6	AX224471 Sequence
38	143.2	9.9	3729	6	AX658011 Sequence
39	143.2	9.9	5426	6	AX224518 Sequence
40	143.2	9.9	5448	6	AR338755 Sequence
41	143.2	9.9	5497	9	AB067488 Homo sapi
42	143.2	9.9	5508	6	AX658014 Sequence
43	143.2	9.9	5511	9	HS0800531
44	143.2	9.9	5532	6	AR338754 Sequence
45	140.4	9.7	1997	10	AF093416 Mus muscu

# ALIGNMENTS

RESULT 1  
BC028587  
LOCUS BC028587 1854 bp mRNA linear PRI 07-OCT-2003  
DEFINITION Homo sapiens NIMA (never in mitosis gene a) - related kinase 11,  
mRNA (CDNA clone MGC:27011 IMAGE:4829802), complete cds.  
ACCESSION BC028587  
VERSION BC028587.2 GI:34191901  
KEYWORDS MGC.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
REFERENCE 1 (bases 1 to 1854)  
AUTHORS Strausberg,R.L., Feingold,E.A., Grouse,L.H., Derge,J.G.,  
Klausner,R.D., Collins,F.S., Wagner,L., Shenmen,C.M., Schuler,G.D.,

Altschul, S.F., Zeeberg, B., Buetow, K.H., Schaefer, C.F., Bhat, N.K., Hopkins, R.F., Jordan, H., Moore, T., Max, S.I., Wang, J., Hsieh, F., Diatchenko, L., Marusina, K., Farmer, A.A., Rubin, G.M., Hong, L., Stapleton, M., Soares, M.B., Bonaldo, M.F., Casavant, T.L., Schetz, T.B., Brownstein, M.J., Usdin, T.B., Toshiyuki, S., Carninci, P., Frange, C., Raha, S., Loquellano, N.A., Peters, G.J., Abramson, R.D., Mullany, S.J., Bosak, S.A., McEwan, P.J., McKernan, K.J., Malek, J.A., Gunaratne, P.H., Richards, S., Worley, K.C., Hale, S.K., Garcia, A.M., Gay, L.J., Hulyk, S.W., Villalón, D.K., Muzny, D.M., Sodergren, E.J., Lu, X., Gibbs, R.A., Fahey, J., Helton, E., Kettman, M., Madan, A., Rodriguez, S., Sanchez, A., Whiting, M., Madan, A., Young, A.C., Shevchenko, Y., Bouffard, G.C., Blakesley, R.W., Touchman, J.W., Green, E.D., Dickson, M.C., Rodriguez, A.C., Grimwood, J., Schmutz, J., Myers, R.M., Butterfield, Y.S., Krzywinski, M.I., Skalska, U., Smalhus, D.E., Schnerch, A., Schein, J.E., Jones, S.J., and Marra, M.A.

Generation and initial analysis of more than 15,000 full-length human and mouse cDNA sequences

Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)

2238257

12477932

2 (bases 1 to 1854)

Strausberg, R.

Direct Submission

Submitted (23-APR-2002) National Institutes of Health, Mammalian Gene Collection (MGC), Cancer Genomics Office, National Cancer Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590, USA

## TITLE

## JOURNAL

## MEDLINE

## PUBMED

## REFERENCE

## AUTHORS

## TITLE

## JOURNAL

## REMARK

## COMMENT

NIH-MGC Project URL: <http://mgc.nci.nih.gov>  
On Aug 25, 2003 this sequence version replaced gi:20306789.

Contact: MGC help desk

Email: [cgaabs@mail.nih.gov](mailto:cgaabs@mail.nih.gov)

Tissue Procurement: Miklos Palkovits, M.D., Ph.D.

cDNA Library Preparation: Michael J. Brownstein (NHGRI) & Shiraki

Toshiyuki and Piero Carninci (RIKEN)

cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)

DNA Sequencing by: Institute for Systems Biology

<http://www.systemsbio.org>

contact: [anadan@systemsbiology.org](mailto:anadan@systemsbiology.org)

Anup Madan, Jessica Fahey, Erin Helton, Mark Kettman, Anuradha

Madan, Stephanie Rodrigues, Amy Sanchez and Michelle Whiting

Clone distribution: MGC clone distribution information can be found

through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>

Series: IRAC Plate: 34 Row: 9 Column: 23

This clone was selected for full length sequencing because it

passed the following selection criteria: matched mRNA gi: 22094142.

## FEATURES

## source

Location/Qualifiers

1..1854

/organism="Homo sapiens"

/mol\_type="mRNA"

/db\_xref="taxon:9606"

/clone="MGC:27011 IMAGE:4829802"

/tissue\_type="testis"

/clone\_lib="NIH\_MGC\_97"

/lab\_host="DH10B"

/note="Vector: pBluescript"

1..1854

/gene="NEK11"

/note="synonym: FLJ23495"

/db\_xref="LocusID:79858"

371..1819

/codon\_start=1

/product="NEK11 protein"

/protein\_id="AAH28587.1"

/db\_xref="GI:20306790"

/db\_xref="LocusID:79858"

/translation="MLKFEAAKCVSGSTAIPTYKTLIARYVLOQLKSGSGFTVY  
LVSDKAKGEELKVLKEISVGEINAPNETVOANLEQLSLKLDHPALVKFHASFVEQD  
NFCITTEYCEGRDLDDKIQEYKQAGKIFPENQIIEWFTQLLLGVDMERILLRDLK  
SKNVLKNNLLKIDFGVSLKMGSCDLATITGTPHYNSPEALKMGSDTKSDIWSL  
ACTLYKMCNNAHAFSGSLSLVILKIVEGDTPLSPERYKELNAMESLNKNPSLRP  
SAIEILKIPYLDEQLNLCRYSENTLEDKLDCKQEAHHIINAMQKRIHLQTLRLS

## gene

## CDS

EVQVQTPPRRMBRLKLOADEKARKLKIVPEEKYEENSKMOELSRNFQOLSVDLVH  
EKTHLKGMEKEQEPGRISCSPQEDERWQGRREESDEPLENLPESQIFPSMDLH  
ELESIVEDATSDLVGHGDCNLISLDEYWNKEK"

455..1231

misc\_feature

/note="S TKC; Region: Serine/Threonine protein kinases,  
catalytic domain"

/db\_xref="CDD:smart00220"

## ORIGIN

Query Match 100.0%; Score 1449; DB 9; Length 1854;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 1449; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ATCTGAAATTCACAGGACGCTAAGTGTGTGAGTGGATCAACAGCCATTTCCTTAT 60  
Dy 371 ATCTGAAATTCACAGGACGCTAAGTGTGTGAGTGGATCAACAGCCATTTCCTTAT 430  
Qy 61 CCAAGACCTTGATTGCAAGAGATACGTCTTCAACAAAACTTGGCAGTGGAGTTT 120  
Dy 431 CCAAGACCTTGATTGCAAGAGATACGTCTTCAACAAAACTTGGCAGTGGAGTTT 490  
Qy 121 GGAAGTGTCTATCTGTTTCAGACAGAAAGCCAAACGAGGAGGAGAAATTAAGTACT 180  
Dy 491 GGAAGTGTCTATCTGTTTCAGACAGAAAGCCAAACGAGGAGGAGAAATTAAGTACT 550  
Qy 181 AAGGAAATATCTTGTGAGAACTAAATCCAAATGAAACTGTACAGGCCAATTTGAAGCC 240  
Dy 551 AAGGAAATATCTTGTGAGAACTAAATCCAAATGAAACTGTACAGGCCAATTTGAAGCC 610  
Qy 241 CAATCTCTTCAAGCTGGACACCCAGCCATTGTCAAGTTCATGCAAGTTTGTGGAG 300  
Dy 611 CAATCTCTTCAAGCTGGACACCCAGCCATTGTCAAGTTCATGCAAGTTTGTGGAG 670  
Qy 301 CAAGATAATTTCTGCATTATCACGAGTACTGTGAGGCGCGAGATCTGGACGATATAAT 360  
Dy 671 CAAGATAATTTCTGCATTATCACGAGTACTGTGAGGCGCGAGATCTGGACGATATAAT 730  
Qy 361 CAGGAAATATAA CAAGCTGGAAAAATCTTTCCGAAAAATCAATAATAGATGTTTATC 420  
Dy 731 CAGGAAATATAA CAAGCTGGAAAAATCTTTCCGAAAAATCAATAATAGATGTTTATC 790  
Qy 421 CAGTGTCTGTGGAGTTGACTACATGATGAGGAGGATACCTTCATCGAGACTTAAAG 480  
Dy 791 CAGTGTCTGTGGAGTTGACTACATGATGAGGAGGATACCTTCATCGAGACTTAAAG 850  
Qy 481 TCAAGAGTATTTCTGAAAAATAATCTCTTAAATTTGAGATTTGGAGTTTTCGA 540  
Dy 851 TCAAGAGTATTTCTGAAAAATAATCTCTTAAATTTGAGATTTGGAGTTTTCGA 910  
Qy 541 CTTCATAGGATCTGTGACCTGGCCACAACTTAACTGGACTCCCATATATAGT 600  
Dy 911 CTTCATAGGATCTGTGACCTGGCCACAACTTAACTGGACTCCCATATATAGT 970  
Qy 601 CCTGAGGCTCTGAAAAACCAAGGCTATGACACAAAGTCGAGACATCTGGTCACTGGCATGC 660  
Dy 971 CCTGAGGCTCTGAAAAACCAAGGCTATGACACAAAGTCGAGACATCTGGTCACTGGCATGC 1030  
Qy 661 ATTTTGTATGAGATGTCTGCATGATCATGATTCGCTGGCTCCCAATTTCTTATCCATT 720  
Dy 1031 ATTTTGTATGAGATGTCTGCATGATCATGATTCGCTGGCTCCCAATTTCTTATCCATT 1090  
Qy 721 GTTTTAAAAATTTGTTGAAGGTGACACACCTTCTCTCCCTGAGAGATATCCAAAGAACTA 780  
Dy 1091 GTTTTAAAAATTTGTTGAAGGTGACACACCTTCTCTCCCTGAGAGATATCCAAAGAACTA 1150  
Qy 781 AATGCCATCATGAAAGCATGTTGAACAGAAATCTTTCATTAAGACCATCTGCTATCGAA 840  
Dy 1151 AATGCCATCATGAAAGCATGTTGAACAGAAATCTTTCATTAAGACCATCTGCTATCGAA 1210  
Qy 841 ATTTTAAAAATTCCTTACCTTGTGATGAGCAGTACAGAACCTAATGTGTAGATATTCAGAA 900  
Dy 1211 ATTTTAAAAATTCCTTACCTTGTGATGAGCAGTACAGAACCTAATGTGTAGATATTCAGAA 1270





Db 1014 GTTTTAAAAATTGTTGAAGGTGACACACCTTCTCTCCCTGAGAGATATCCAAAGAACTA 1073  
QY 781 AATGCCATCATGAAAGCATGTTGAACAGAAATCCTTCAATTAAGACCATCTGCTATCGAA 840  
Db 1074 AATGCCATCATGAAAGCATGTTGAACAGAAATCCTTCAATTAAGACCATCTGCTATCGAA 1133  
QY 841 ATTTTAAAAATCCCTTACCTTGTAGTACGAGCTACAGAACTTAATGTGTAGATATTAGAA 900  
Db 1134 ATTTTAAAAATCCCTTACCTTGTAGTACGAGCTACAGAACTTAATGTGTAGATATTAGAA 1193  
QY 901 ATGACTCTGGAAGACAAAATTTGGATTGTGAGAGGAGCTGCTCATATAATTAATGCC 960  
Db 1194 ATGACTCTGGAAGACAAAATTTGGATTGTGAGAGGAGCTGCTCATATAATTAATGCC 1253  
QY 961 ATGCAAAAAGGATCCACCTGACAGCTCTGAGGGCACTGTGAGAGTACAGAAATGAGC 1020  
Db 1254 ATGCAAAAAGGATCCACCTGACAGCTCTGAGGGCACTGTGAGAGTACAGAAATGAGC 1313  
QY 1021 CCAAGGAAAGGATCGGGCTGAGAAAGCTCCAGGGCTGTGATGAGAAAGCCAGGAAGCTG 1080  
Db 1314 CCAAGGAAAGGATCGGGCTGAGAAAGCTCCAGGGCTGTGATGAGAAAGCCAGGAAGCTG 1373  
QY 1081 AAAAAAGNTGTGAGAAAGAAATATGAAGAAATAGCAACGAATGCAAGAAATTTGAGATCT 1140  
Db 1374 AAAAAAGNTGTGAGAAAGAAATATGAAGAAATAGCAACGAATGCAAGAAATTTGAGATCT 1433  
QY 1141 CGAACTTTTCAGCAGCTGAGTGTGTGATGTACTCCATGAAAAACACATTTAAAGGAATG 1200  
Db 1434 CGAACTTTTCAGCAGCTGAGTGTGTGATGTACTCCATGAAAAACACATTTAAAGGAATG 1493  
QY 1201 GAAGAAAGGAGAGCAACTGAGGAAGACATTTCTGTTTCCACCCAGCAGAGATGAA 1260  
Db 1494 GAAGAAAGGAGAGCAACTGAGGAAGACATTTCTGTTTCCACCCAGCAGAGATGAA 1553  
QY 1261 GAGAGTGGCAAGCAGGAGAGGAATCTGTGAACCAACTTTAGAGAACTTGCTGAG 1320  
Db 1554 GAGAGTGGCAAGCAGGAGAGGAATCTGTGAACCAACTTTAGAGAACTTGCTGAG 1613  
QY 1321 TCTCAGCTTATCTCTCCATGAGACCTCCAGAACTTGAATCAATTTGATAGAGATGCCACA 1380  
Db 1614 TCTCAGCTTATCTCTCCATGAGACCTCCAGAACTTGAATCAATTTGATAGAGATGCCACA 1673  
QY 1381 TCTGACCTTGGATACCATGAGAGCTGTAATCTAATTCA 1419  
Db 1674 TCTGACCTTGGATACCATGAGAGCTGTAATCTAATTCA 1712

RESULT 3  
AX166555 LOCUS 1938 bp DNA linear PAT 22-JUN-2001  
DEFINITION Sequence 46 from Patent WO0138503.  
ACCESSION AX166555  
VERSION AX166555.1 GI:14546900  
KEYWORDS Homo sapiens (human)  
SOURCE Homo sapiens  
ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
REFERENCE 1 Plowman,G.D., Whyte,D., Manning,G.S., Sudarsanam,S.S., Martinez,R.,  
AUTHORS Flanagan,P. and Clary,D.S.  
TITLE Novel human protein kinases and protein kinase-like enzymes  
JOURNAL Patent: WO 0138503-A 46 31-MAY-2001;  
Sugen, Inc. (US)  
FEATURES  
source Location/Qualifiers  
1..1938  
/organism="Homo sapiens"  
/mol\_type="unassigned DNA"  
/db\_xref="taxon:9606"

ORIGIN

Query Match

96.5%; Score 1399; DB 6; Length 1938;

Best Local Similarity 100.0%; Pred. No. 0;  
Matches 1399; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 ATGCTGAAATTCGAAGAGGAGCTAAGTGTGAGTGGATCAACAGCACTTCCACTTAT 60  
Db 1 ATGCTGAAATTCGAAGAGGAGCTAAGTGTGAGTGGATCAACAGCACTTCCACTTAT 60  
QY 61 CCAAGACCTTGATTCGAAGAGATACGTCCTTCAACAAAACTTGGCAGTGAAGTTT 120  
Db 61 CCAAGACCTTGATTCGAAGAGATACGTCCTTCAACAAAACTTGGCAGTGAAGTTT 120  
QY 121 GGAATCTCTATCTGTTTTCAGACACAGAAAGCCAAACGAGGAGAGAAATTAAGAGTACTT 180  
Db 121 GGAATCTCTATCTGTTTTCAGACACAGAAAGCCAAACGAGGAGAGAAATTAAGAGTACTT 180  
QY 181 AAGGAAATATCTGTTGAGAACTAAATCCAAATCTACAGGCCAAATTTGGAAGCC 240  
Db 181 AAGGAAATATCTGTTGAGAACTAAATCCAAATCTACAGGCCAAATTTGGAAGCC 240  
QY 241 CAATCTCTTCCAGCTGGACCCAGCCCACTTGTCAAGTTCCTCATGCACTTTTGGAG 300  
Db 241 CAATCTCTTCCAGCTGGACCCAGCCCACTTGTCAAGTTCCTCATGCACTTTTGGAG 300  
QY 301 CAAGATAATTTCTGCACTTATCAAGGAGTACTGTGAGGGCCGAGATCTGGACGATAAAT 360  
Db 301 CAAGATAATTTCTGCACTTATCAAGGAGTACTGTGAGGGCCGAGATCTGGACGATAAAT 360  
QY 361 CAGGAATATAAACAAGCTGGAAAAATCTTTCCAGAAAAATCAAAATATAGATGGTTTATC 420  
Db 361 CAGGAATATAAACAAGCTGGAAAAATCTTTCCAGAAAAATCAAAATATAGATGGTTTATC 420  
QY 421 CAGCTGCTGCTGGAGTTGACTACATCATGAGAGGAGGATCTTCATCGAGACTTAAAG 480  
Db 421 CAGCTGCTGCTGGAGTTGACTACATCATGAGAGGAGGATCTTCATCGAGACTTAAAG 480  
QY 481 TCAAGAAATGTATTTCTGAAAAATTAATCTCTTAAAAATTTGGAGTTTTCGGA 540  
Db 481 TCAAGAAATGTATTTCTGAAAAATTAATCTCTTAAAAATTTGGAGTTTTCGGA 540  
QY 541 CTTCTAATGGATCTGCTGACCTGGCCACAACTTTAACTGGAACCTCCCACTTATATGAT 600  
Db 541 CTTCTAATGGATCTGCTGACCTGGCCACAACTTTAACTGGAACCTCCCACTTATATGAT 600  
QY 601 CCTGAGGCTCTGAAACCAAGGCTATGACACAAAGTCGACATCTGGTCACTGGCATGC 660  
Db 601 CCTGAGGCTCTGAAACCAAGGCTATGACACAAAGTCGACATCTGGTCACTGGCATGC 660  
QY 661 ATTTTGTATGAGATGCTGCTCATGAATCATGCAATTCGCTGGCTCCCAATTTCTTCCATT 720  
Db 661 ATTTTGTATGAGATGCTGCTCATGAATCATGCAATTCGCTGGCTCCCAATTTCTTCCATT 720  
QY 721 GTTTTAAAAATTTGTTGAAGGTGACACACCTTCTCTCCCTGAGAGATATCCAAAGAACTA 780  
Db 721 GTTTTAAAAATTTGTTGAAGGTGACACACCTTCTCTCCCTGAGAGATATCCAAAGAACTA 780  
QY 781 AATGCCATCATGAAAGCATGTTGAACAGAAATCCTTCAATTAAGACCATCTGCTATCGAA 840  
Db 781 AATGCCATCATGAAAGCATGTTGAACAGAAATCCTTCAATTAAGACCATCTGCTATCGAA 840  
QY 841 ATTTTAAAAATCCCTTACCTTGTAGTACGAGCTACAGAACTTAATGTGTAGATATTAGAA 900  
Db 841 ATTTTAAAAATCCCTTACCTTGTAGTACGAGCTACAGAACTTAATGTGTAGATATTAGAA 900  
QY 901 ATGACTCTGGAAGACAAAATTTGGATTGTGAGAGGAGCTGCTCATATAATTAATGCC 960  
Db 901 ATGACTCTGGAAGACAAAATTTGGATTGTGAGAGGAGCTGCTCATATAATTAATGCC 960  
QY 961 ATGCAAAAAGGATCCACCTGACAGCTCTGAGGGCACTGTGAGAGTACAGAAATGAGC 1020  
Db 961 ATGCAAAAAGGATCCACCTGACAGCTCTGAGGGCACTGTGAGAGTACAGAAATGAGC 1020  
QY 1021 CCAAGGAAAGGATCGGGCTGAGAAAGCTCCAGGGCTGTGATGAGAAAGCCAGGAAGCTG 1080

Db 1021 CCAAGAGAAAGGATCGGCTGAGGAAGCTCCAGGCGGCTGATGAGAAAAGCCAGGAAGCTG 1080  
Qy 1081 AAAAGATTGTGGAAGAAAATATCAAGAAAATAGCAAGAAATCAAGAAATGAGATCT 1140  
Db 1081 AAAAGATTGTGGAAGAAAATATGAGAAAATAGCAAGAAATGAGATCT 1140  
Qy 1141 CGAACTTTCAGCAGCTGAGTGTGATGTACTCCATGAAAACACATTTAAAGGAATG 1200  
Db 1141 CGAACTTTCAGCAGCTGAGTGTGATGTACTCCATGAAAACACATTTAAAGGAATG 1200  
Qy 1201 GAAGAAAGGAGGACCACTGAGGAGAGACTTCTGTTTCAACCCAGGAGCGAGGATGAA 1260  
Db 1201 GAAGAAAGGAGGAGCACTGAGGAGAGACTTCTGTTTCAACCCAGGAGCGAGGATGAA 1260  
Qy 1261 GAGAGTGGCAAGGAGGAGGAGGAGTCTGATGAACCAACTTTAGAGAACTGCGCTGAG 1320  
Db 1261 GAGAGTGGCAAGGAGGAGGAGGAGTCTGATGAACCAACTTTAGAGAACTGCGCTGAG 1320  
Qy 1321 TCTGAGCTTATTCCTTCCATGAGACCTCCAGAACTTGAATCAATTTGAGAGATGCCACA 1380  
Db 1321 TCTGAGCTTATTCCTTCCATGAGACCTCCAGAACTTGAATCAATTTGAGAGATGCCACA 1380  
Qy 1381 TCTGAGCTTATTCCTTCCATGAGACCTCCAGAACTTGAATCAATTTGAGAGATGCCACA 1380  
Db 1381 TCTGAGCTTATTCCTTCCATGAGACCTCCAGAACTTGAATCAATTTGAGAGATGCCACA 1380

RESULT 4  
AX746174  
LOCUS AX746174 2552 bp DNA linear PAT 13-JUN-2003  
DEFINITION Sequence 25 from Patent WO2008399.  
ACCESSION AX746174  
VERSION AX746174.1 GI:31746160  
KEYWORDS  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
REFERENCE 1  
AUTHORS Yue,H., Khan,F.A., Gururajan,R., Hafalia,A.J., Chawla,N.K.,  
Arvizu,C.S., Ramkumar,J., Gandhi,A.R., Policky,J.L., Baughn,M.R.,  
Tribouley,C.M., Bandman,O., Nguyen,D.B., Lu,Y., Burford,N., Lal,P.,  
Ding,L., Yao,M.G., Elliott,V.S., Recipon,S.A., Kearney,L., Lu,D.A.,  
Greenwald,S.R., Tang,Y.T., Xu,Y., Walsh,R.T., Gietzen,K.J.,  
Yang,J., Jackson,J.L. and Thornton,M.  
Human kinases  
TITLE Human kinases  
JOURNAL Patent: WO 0208399-A 25 31-JAN-2002;  
Incyte Genomics, Inc. (US) ; Thornton, Michael (US)  
FEATURES  
Location/Qualifiers  
1..2552  
/organism="Homo sapiens"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:9606"  
/note="Incyte ID No: 7477585CB1"

ORIGIN  
Query Match 96.5%; Score 1399; DB 6; Length 2552;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 1399; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Qy 1 ATGCTGAATTCAGAGCAGCTAAGTGTGTGAGTGGATCAACGCCATTTCCACTTAT 60  
Db 278 ATGCTGAATTCAGAGCAGCTAAGTGTGTGAGTGGATCAACGCCATTTCCACTTAT 337  
Qy 61 CCAAGACCTTGATTCGAAGAGATACGTGCTTCAACAAAACCTTGGCAGTGGAGTTT 120  
Db 338 CCAAGACCTTGATTCGAAGAGATACGTGCTTCAACAAAACCTTGGCAGTGGAGTTT 397  
Qy 121 GGAAGTGTCTATCTGTTTTCAGACAAAGAACCAACGAGGAGAGAAATTAAGTACT 180  
Db 398 GGAAGTGTCTATCTGTTTTCAGACAAAGAACCAACGAGGAGAGAAATTAAGTACT 457  
Qy 181 AGGAATATCTGTTGGAGAACTAATCAATGAACTGTACAGGCCAATTTGGAAGCC 240

Db 458 AAGGAATATCTGTTGGAGAACTAAATCCAAATGAAATCTGACAGGCCAATTTGGAAGCC 517  
Qy 241 CAATCTCTCTCCAAAGCTGGACCAACCCAGCCATTTGCAAGTTCCATGCAAGCTTTTGTGGAG 300  
Db 518 CAATCTCTCTCCAAAGCTGGACCAACCCAGCCATTTGCAAGTTCCATGCAAGCTTTTGTGGAG 577  
Qy 301 CAAGATAAATTTGCAATATCACGGAGTACTGTGAGGCGCGAGATCTGGACGATATAAATT 360  
Db 578 CAAGATAAATTTGCAATATCACGGAGTACTGTGAGGCGCGAGATCTGGACGATATAAATT 637  
Qy 361 CAGGAATATAAACAAGCTGGAAAATCTTCCAGAAAATCABAATATAGATGTTTATC 420  
Db 638 CAGGAATATAAACAAGCTGGAAAATCTTCCAGAAAATCABAATATAGATGTTTATC 697  
Qy 421 CAGCTGCTGCTGGAGTTGACTACATGATGAGAGGAGTACTTTCATCGAGACTTTAAAG 480  
Db 698 CAGCTGCTGCTGGAGTTGACTACATGATGAGAGGAGTACTTTCATCGAGACTTTAAAG 757  
Qy 481 TCAAGAATATGATTTCTGAAAATATCTCTTTAAAATTTGGAGATTTTGGAGTTTCTCGA 540  
Db 758 TCAAGAATATGATTTCTGAAAATATCTCTTTAAAATTTGGAGATTTTGGAGTTTCTCGA 817  
Qy 541 CTTCTAATGGATCTCTGACCTGGCCACAACTTTAACTGAAAACCTCCCAATTATATGAGT 600  
Db 818 CTTCTAATGGATCTCTGACCTGGCCACAACTTTAACTGAAAACCTCCCAATTATATGAGT 877  
Qy 601 CTTGAGGCTCTGAAAACCAAGGCTATGACAAAAGTCGACATCTGGTCACTTGGCATGC 660  
Db 878 CTTGAGGCTCTGAAAACCAAGGCTATGACAAAAGTCGACATCTGGTCACTTGGCATGC 937  
Qy 661 ATTTTGTATGAGATGTGCTGCATGAATCATGCAATTCGCTGGCTCCAAATTTCTTATCCATT 720  
Db 938 ATTTTGTATGAGATGTGCTGCATGAATCATGCAATTCGCTGGCTCCAAATTTCTTATCCATT 997  
Qy 721 GTTTTAAAAATTTGTGAAGTGCACACCTTCTCTCCCTGAGAGATATCAAAAGAACTA 780  
Db 998 GTTTTAAAAATTTGTGAAGTGCACACCTTCTCTCCCTGAGAGATATCAAAAGAACTA 1057  
Qy 781 AATGCCATCATGGAAGCATGTTGAAACAAGATCCTTCAATAGACCATCTGCTATCGAA 840  
Db 1058 AATGCCATCATGGAAGCATGTTGAAACAAGATCCTTCAATAGACCATCTGCTATCGAA 1117  
Qy 841 ATTTTAAAAATCCCTTACCTTGTATGACAGCTACAGAACCTAATGTGTAGATATTCAGAA 900  
Db 1118 ATTTTAAAAATCCCTTACCTTGTATGACAGCTACAGAACCTAATGTGTAGATATTCAGAA 1177  
Qy 901 ATGACTCTGGAAGACAAAATTTGGATGTGACAGAGGAGCTGCTCATATATTAATGCC 960  
Db 1178 ATGACTCTGGAAGACAAAATTTGGATGTGACAGAGGAGCTGCTCATATTAATGCC 1237  
Qy 961 ATGCAAAAAGGATCCACCTGACACTCTGAGGCGACTGTGAGGAGTACAGAAAATGACG 1020  
Db 1238 ATGCAAAAAGGATCCACCTGACACTCTGAGGCGACTGTGAGGAGTACAGAAAATGACG 1297  
Qy 1021 CCAAGAGAAAGGATCGGCTGAGAAAGCTCCAGGCGCTGATGAGAAAGCCAGGAAGCTG 1080  
Db 1298 CCAAGAGAAAGGATCGGCTGAGAAAGCTCCAGGCGCTGATGAGAAAGCCAGGAAGCTG 1357  
Qy 1081 AAAAGATTGTGGAAGAAAATATGAGAAAATAGCAAGAAATGCAAGAAATTTGAGATCT 1140  
Db 1358 AAAAGATTGTGGAAGAAAATATGAGAAAATAGCAAGAAATGCAAGAAATTTGAGATCT 1417  
Qy 1141 CGGAACCTTTCAGCAGCTGAGTGTGATGTACTCCATGAAAACACATTTAAAGGAATG 1200  
Db 1418 CGGAACCTTTCAGCAGCTGAGTGTGATGTACTCCATGAAAACACATTTAAAGGAATG 1477  
Qy 1201 GAAAGAAAGGAGGAGCAACCTGAGGAGAGACTTCTGTTTCAACCCAGGAGGAGGATGAA 1260  
Db 1478 GAAAGAAAGGAGGAGCAACCTGAGGAGAGACTTCTGTTTCAACCCAGGAGGAGGATGAA 1537  
Qy 1261 GAGAGTGGCAAGGAGGAGGAGGATCTGATGAACCAACTTTAGAGAACTTTCGCTGAG 1320

```

Db      1538 GAGAGTGGCAAGGAGGAGGAATCTGATGAACCAACTTTAGAGAACCTGCTGAG 1597
QY      1321 TCTCAGCCTATTCCTTCCATGGACCTCCAGAACTTGAATCAATTTGTAGAGGATGCCACA 1380
Db      1598 TCTCAGCCTATTCCTTCCATGGACCTCCAGAACTTGAATCAATTTGTAGAGGATGCCACA 1657
QY      1381 TCTGACCTTGGATACCATG 1399
Db      1658 TCTGACCTTGGATACCATG 1676

RESULT 5
HSR804785
LOCUS      2768 bp mRNA linear PRI 13-MAY-2003
DEFINITION Homo sapiens mRNA; cDNA DKF2p686L0617 (from clone DKF2p686L0617).
ACCESSION AL633472
VERSION    AL633472.1 GI:21734115
KEYWORDS
SOURCE     Homo sapiens (human)
ORGANISM   Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 2768)
AUTHORS    Bahr,A., Lauber,J., Mewes,H.W., Weil,B., Amid,C., Osanger,A.,
            Fobo,G., Han,M. and Wiemann,S.
TITLE      Direct Submission
JOURNAL    Submitted (13-MAY-2003) MIPS, Ingolstaedter Landstr.1, D-85764
            Neuherberg, GERMANY
COMMENT     Clone from S. Wiemann, Molecular Genome Analysis, German Cancer
            Research Center (DKFZ); Email s.wiemann@dkfz-heidelberg.de;
            sequenced by Qiagen (Hilden/Germany) within the cDNA sequencing
            consortium of the German Genome Project.
            This clone (DKF2p686L0617) is available at the RZPD in Berlin.
            Please contact the RZPD: Ressourcenzentrum, Heubnerweg 6, 14059
            Berlin-Charlottenburg, GERMANY; Email: clone@rzpd.de Further
            information about the clone and the sequencing project is available
            at http://mips.gsf.de/proj/cDNA/.
FEATURES
     source              1..2768
                        /organism="Homo sapiens"
                        /mol_type="mRNA"
                        /db_xref="taxon:9606"
                        /clone="DKF2p686L0617"
                        /tissue_type="human cervix"
                        /clone_lib="686 (synonym: hicc3). Vector pSport1_Sfi; host
                        DH10B; sites SfiIA + SfiIB"
                        /dev_stage="adult"
     polyA_signal        2726..2731
     polyA_site          2753
     ORIGIN
Query Match      96.5%; Score 1399; DB 9; Length 2768;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1399; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 ATGCTGAATTCACAGCAGCAGCTAAGTCTGTGAGTGGATCAACAGCCATTTCCACTTAT 60
Db      225 ATGCTGAATTCACAGCAGCAGCTAAGTCTGTGAGTGGATCAACAGCCATTTCCACTTAT 284
QY      61 CCAAGACCTTGTATTCGAAGAAGATACGTGCTTCAACAAAACTTGGCAGTGGAAAGTTT 120
Db      285 CCAAGACCTTGTATTCGAAGAAGATACGTGCTTCAACAAAACTTGGCAGTGGAAAGTTT 344
QY      121 GGAACTGTCTATCTGTTTTCAGACAAAGAGCCAAACGAGGAGAGGAATTAAGTACTT 180
Db      345 GGAACTGTCTATCTGTTTTCAGACAAAGAGCCAAACGAGGAGAGGAATTAAGTACTT 404
QY      181 AAGGAATATCTGTTGGAGAACTAATCCAAATGAACCTGTACAGGCCAAATTTGGAAGCC 240
Db      405 AAGGAATATCTGTTGGAGAACTAATCCAAATGAACCTGTACAGGCCAAATTTGGAAGCC 464
QY      241 CAACTCCTTCCAAAGCTGAGCACCCAGCCATGTGCAAGTTCATGCAAGTTTGTGGAG 300

```

```

Db      465 CAATCTCTTCCAAAGCTGGACCAACCAGCCATTTGTCAAGTTCCATGCAAGTTTGTGGAG 524
QY      301 CAAGATAATTTCTGCATTATCAAGAGTACTGTGAGGGCCGAGACTGTGGACGATAAAATT 360
Db      525 CAAGATAATTTCTGCATTATCAAGAGTACTGTGAGGGCCGAGACTGTGGACGATAAAATT 584
QY      361 CAGGAATATAAACAAGCTGGAAAATCTTTCCAGAAATCAAAATTAATAGAAATGTTTATC 420
Db      585 CAGGAATATAAACAAGCTGGAAAATCTTTCCAGAAATCAAAATTAATAGAAATGTTTATC 644
QY      421 CAGCTGCTGCTGGAGTTGACTACATGTCATGAGGAGGAGATCTTTCATCGAGACTTTAAAG 480
Db      645 CAGCTGCTGCTGGAGTTGACTACATGTCATGAGGAGGAGATCTTTCATCGAGACTTTAAAG 704
QY      481 TCAAGAAGATGTTTCTGAAAAATATCTCTTAAATTTGGAGATTTTGGAGTTTCTCGA 540
Db      705 TCAAGAAGATGTTTCTGAAAAATATCTCTTAAATTTGGAGATTTTGGAGTTTCTCGA 764
QY      541 CTTCTAATGGGATCCTGTGACCTGGCCACAACTTTAACTGGAACTCCCATATATAGT 600
Db      765 CTTCTAATGGGATCCTGTGACCTGGCCACAACTTTAACTGGAACTCCCATATATAGT 824
QY      601 CTTGAGGCTCTGAAAACACCAAGGCTATGACAAAGTGGACATCTGGTCACTGGGATGC 660
Db      825 CTTGAGGCTCTGAAAACACCAAGGCTATGACAAAGTGGACATCTGGTCACTGGGATGC 884
QY      661 ATTTTGTATGAGATGTCTGTCATGATCATGATTCATGCTGGCTCCAAATTTCTTATCAT 720
Db      885 ATTTTGTATGAGATGTCTGTCATGATCATGATTCATGCTGGCTCCAAATTTCTTATCAT 944
QY      721 GTTTTAAAAATTTGTTGAAGTGACACACCTTCTCTCCCTGAGAGATATCCAAAAGAACTA 780
Db      945 GTTTTAAAAATTTGTTGAAGTGACACACCTTCTCTCCCTGAGAGATATCCAAAAGAACTA 1004
QY      781 AATGCCATCATGAAAGCATGTTTGAACAGAACTCTTCAATTAAGACCATCTGCTATCGAA 840
Db      1005 AATGCCATCATGAAAGCATGTTTGAACAGAACTCTTCAATTAAGACCATCTGCTATCGAA 1064
QY      841 ATTTTAAAAATTCCTTTACCTTGTATGAGCAGCTACAGAACCTTAATGTGTAGATATTCAGAA 900
Db      1065 ATTTTAAAAATTCCTTTACCTTGTATGAGCAGCTACAGAACCTTAATGTGTAGATATTCAGAA 1124
QY      901 ATGACTCTGGAAGACAAAAATTTGGATTTGTCAAGAGAGGCTCTCATATATTAATGCC 960
Db      1125 ATGACTCTGGAAGACAAAAATTTGGATTTGTCAAGAGAGGCTCTCATATATAATATGCC 1184
QY      961 ATGCAAAAAAGGATCCACCTGACAGCTCTGAGGGCAGCTGTCAAGATACAGAAAAATGACG 1020
Db      1185 ATGCAAAAAAGGATCCACCTGACAGCTCTGAGGGCAGCTGTCAAGATACAGAAAAATGACG 1244
QY      1021 CCAAGAGAAAAGGATGCGGCTGAGGAAGTCTCAGCGGCTGTATGAGAAAAGCCAGGAAGCTG 1080
Db      1245 CCAAGAGAAAAGGATGCGGCTGAGGAAGTCTCAGCGGCTGTATGAGAAAAGCCAGGAAGCTG 1304
QY      1081 AAAAAAGTTCTGGAAGAAAAATATGAAGAAAAATAGCAAAACGAATGCAAGAAATTCAGATCT 1140
Db      1305 AAAAAAGTTCTGGAAGAAAAATATGAAGAAAAATAGCAAGAAATTCAGATTCAGATCT 1364
QY      1141 CGGAACCTTTCAGCAGCTGAGTGTGATGTAATCCATGAAAAAACACATTTAAAGGAATG 1200
Db      1365 CGGAACCTTTCAGCAGCTGAGTGTGATGTAATCCATGAAAAAACACATTTAAAGGAATG 1424
QY      1201 GAAGAAAGCAGGAGCAACCTGAGGAGAGACTTTCTTTGTTTCACTCCAGGAGCAGGAGTGA 1260
Db      1425 GAAGAAAGCAGGAGCAACCTGAGGAGAGACTTTCTTTGTTTCACTCCAGGAGCAGGAGTGA 1484
QY      1261 GAGAGGTGGCAAGCGAGGGAAGAGGAATCTGTAAACCACTTTAGAGAACCTGCTCGTAG 1320
Db      1485 GAGAGGTGGCAAGCGAGGGAAGAGGAATCTGTAAACCACTTTAGAGAACCTGCTCGTAG 1544
QY      1321 TCTCAGCTATTTCTTCCATGAGCTCCAGGACTTTGAATCAATTTAGAGGATGCCACA 1380
Db      1545 TCTCAGCTATTTCTTCCATGAGCTCCAGGACTTTGAATCAATTTAGAGGATGCCACA 1604

```







955 ATTTTATATGAGATGCTGTCATGAATCATGCAATTCGCTGGCTCCOAAATTTCTTTATCCATT 1014  
 721 GTTTTAAAAAATTTGTTGAAGGTGACACACCTTCTCTCCCTGAGAGATATCAAAAGAACTA 780  
 1015 GTTTTAAAAAATTTGTTGAAGGTGACACACCTTCTCTCCCTGAGAGATATCAAAAGAACTA 1074  
 781 AATGCCATCATGGAAGCATGTTGAACAAGATCCCTTCATTAAGACCATCTGCTATCGAA 840  
 1075 AATACCATCATGGAAGCATGTTGAACAAGATCCCTTCATTAAGACCATCTGCTATCGAA 1134  
 841 ATTTTAAAAAATCCCTTACCTTCTGATGAGCAGCTACAGAACTATGTGTAGATATTCAGAA 900  
 1135 ATTTTAAAGATCCCTTACATTTGATGAGCAGCTACAGCACTATGTGTAGCATTCAGAA 1194  
 901 ATGACTCTGGAAGACAAAATTTGGATTGTCAAGAGAGCGTCTCTCATATAATATGCC 960  
 1195 ATGACTCTGGAAGACAAAATTTGGATTGTCAAGAGAGCGTCTCTCATATAATATGCC 1254  
 961 ATGCAAAAAGGATCACCTGTCAGACTCTTGAGGGCACTGTCAAGACTACAGAAAATGACG 1020  
 1255 ATGCAAAAAGGATCACCTGTCAGACTCTTGAGGGCCCTGTCAAGACTACAGAAAATGACT 1314  
 1021 CCAAGAAAAGGATCGGCTGAGGAAGCTCCAGGCGCTGATGAGAAAACCAAGAAAGCTG 1080  
 1315 CCAAGAAAAGGATGAGGCTGAGAAGCTCCAGGCACTGATGAGAGCCAGGAAGCTG 1374  
 1081 AAAAAAGATTGCGAAGAAAAATATGAAGAAAATAGCAAACGATGCAAGAAATTTGATCT 1140  
 1375 AAAAAAGATTGCGAAGAAAAATATGAAGAAAATAGCAAACGATGCAAGAAATTTGATCT 1434  
 1141 CGGAATTTACGACGTGAGTGTGATGTACTCCATGAAAAAACACATTTAAAGGAATG 1200  
 1435 CGGAATTTACGACGTGAGTGTGATGTACTCCATGAAAAAACACATTTAAAGGAATG 1494  
 1201 GAAGAAAAGGAGGACCACTCAGGGAAGCACTTCTGTCTACCCGAGCAGGAGATGAA 1260  
 1495 GAAGAAAAGGAGGAGCAACTCAGGGAAGCACTTCTGTCTACCCGAGCAGGAGATGAA 1554  
 1261 GAGAGGTGGCAAGGAGGAGGAAGAAATCTGATGAACCAACTTTAGAGAACCCTGCCTGAG 1320  
 1555 GAGAGGTGGCAAGGAGGAGGAATTTGATGAACCGACTTTAGAGAACCCTGCTGAA 1614  
 1321 TCTGAGCCTATTCTTCATGACCTCCAGCAACTTGAATCAATTTAGAGAGATGCCACA 1380  
 1615 CTTGAGCCTATTCTTCATGACCTCCAGCAACTTGAATCAATTTAGAGAGATGCCACA 1674  
 1381 TCTGACCTTGGATACCATG 1399  
 1675 TCTGACCTTGGATATCATG 1693  
 RESULT 8  
 AK027148  
 LOCUS  
 DEFINITION Homo sapiens cDNA: FLJ23495 fis, clone LNG02228.  
 ACCESSION AK027148  
 VERSION AK027148.1 GI:10440201  
 KEYWORDS oligo capping; fis (full insert sequence).  
 SOURCE Homo sapiens (human)  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 1  
 REFERENCE  
 AUTHORS Kawakami, T., Noguchi, S., Itoh, T., Shigeta, K., Senba, T.,  
 Matsumura, K., Nakajima, Y., Mizuno, T., Morinaga, M., Tanigami, A.,  
 Fujiwara, T., Ono, T., Yamada, K., Fujii, Y., Ozaki, K., Hirao, M.,  
 Ohnori, Y., Oka, T., Suzuki, Y., Obayashi, M., Nishi, T., Shibahara, T.,  
 Tanaka, T., Nakamura, Y., Isogai, T. and Sugano, S.  
 NEDO human cDNA sequencing project  
 Unpublished  
 JOURNAL  
 REFERENCE 2 (bases 1 to 2049)  
 AUTHORS Sugano, S., Suzuki, Y., Oka, T., Obayashi, M., Nishi, T., Isogai, T.,

Snibahara, T., Tanaka, T. and Nakamura, Y.  
Direct Submission  
Submitted (29-AUG-2000) Sumio Sugano, Institute of Medical Science,  
University of Tokyo, Laboratory of Genome Structure Analysis, Human  
Genome Center, Shirokane-dai, 4-6-1, Minato-ku, Tokyo 108-8639,  
Japan (E-mail: flicdn@ims.u-tokyo.ac.jp, Tel: 81-3-5449-5286,  
Fax: 81-3-5449-5416)  
NEDO human cDNA sequencing project supported by Ministry of  
International Trade and Industry of Japan; cDNA full insert  
sequencing: Research Association for Biotechnology; cDNA library  
construction, 5'- & 3'-end one pass sequencing: Department of  
Virology and Human Genome Center, Institute of Medical Science,  
University of Tokyo (partly supported by Science and Technology  
Agency).

```

FEATURES
source
Location/Qualifiers
1..2049
/organism="Homo sapiens"
/mol_type="rRNA"
/db_xref="taxon:9606"
/clone="LNG02228"
/tissue_type="human lung"
/clone_lib="LNG"
/note="cloning vector pME18SF13"
66..1457
/note="unnamed protein product"
/codon_start=1
/protein_id="BAB15672.1"
/db_xref="GI:10440202"
/translation="MGSCDGLATTLTGTPHYVSPEALCHQGYDTSKDIWLSLACILVEM
CMNHAFAAGSNFLISVLKIVTEGDTSLPERYPVKELNAIMSNKNPSLRPSAIEILKI
PYDLQOLNLCRYSEMTEDKNDLDCQEAHHINAMQKRIHLQTLRALSEVQKWTTP
EYMRUKQCADEAKRKLUKIVERKIESNRMQBELSRNFQQSLPVDVHLKTHLKG
EKEBQPGRLSCSQPQDEBERQDEEESDEPTLENLPESQPIPSMDLHLELESIVD
ATSDIYHEIIPEDPLVAEYADAFDSYCSVEDEEEIALERPEKSTRNKGSPAYR
INQDQTEALCARLGNLGVGTSLDTKITITMAEDMGPGPPIFNSVMARTKMKRMRES
AMKQIGTEVFSEVNYLKRASHQNASEAIRIECKVVPQASDCFVDQLLYFFEBQLL
ITMKEPTLONHL"

```

```

ORIGIN
Query Match 63.4%; Score 918; DB 9; Length 2049;
Best Local Similarity 100.0%; Pred. No. 3e-222;
Matches 918; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 482 CAAGAAGTATTTCTGAAAAATAATCTCCCTTAAATTTGGAGATTTTCGGAGTTTCTCGAC 541
DB 1 CAAGAAGTATTTCTGAAAAATAATCTCCCTTAAATTTGGAGATTTTCGGAGTTTCTCGAC 60

QY 542 TTCTAATGGGATCTGTGACCTGGGCCACAACATTTAACTGGAACCTCCCATTATATGAGTC 601
DB 61 TTCTAATGGGATCTGTGACCTGGGCCACAACATTTAACTGGAACCTCCCATTATATGAGTC 120

QY 602 CTGAGGCTCTGAACACCAACGAGCTATGACACAAAGTCGGACATCTGTCACCTGCAGTCGA 661
DB 121 CTGAGGCTCTGAACACCAACGAGCTATGACACAAAGTCGGACATCTGTCACCTGCAGTCGA 180

QY 662 TTTTCTATGAGATGTGCTGCGATGAATCATGANTCGCTGGCTCCAAATTTCTATCCATTG 721
DB 181 TTTTCTATGAGATGTGCTGCGATGAATCATGANTCGCTGGCTCCAAATTTCTATCCATTG 240

QY 722 TTTTAAAAATTTGAAGGTCGACACACCTTCTCCCTCGAGAGATATCCAAAAAGAACTAA 781
DB 241 TTTTAAAAATTTGAAGGTCGACACACCTTCTCCCTCGAGAGATATCCAAAAAGAACTAA 300

QY 782 ATGCCATCATGGAAGAGCATGTTGAACAGAGATCTTTCATTAAGACCACTCTGCTATCGAAA 841
DB 301 ATGCCATCATGGAAGAGCATGTTGAACAGAGATCTTTCATTAAGACCACTCTGCTATCGAAA 360

QY 842 TTTTAAAAATCCCTTACCTTGATGAGCAGCTACAGAACCTTAATGTGTAGATATTCAGAAA 901
DB 361 TTTTAAAAATCCCTTACCTTGATGAGCAGCTACAGAACCTTAATGTGTAGATATTCAGAAA 420

QY 902 TGACTCTGGAAGACAAAAATTTGGATTGTCGAGAGGAGGCTGCTCATATAATTTAATGCCA 961
DB 961

```





QY 601 CCTGAGGCTCTGAACACCAAGGCTATGACACAAAGTCGACATCTGTCATCTGGCATGC 660  
Db 766 CTTGAGGCTTTGAACACCAAGGCTATGACACAAAGTCGACATCTGTCATCTGGCATGC 825  
QY 661 ATTTTGTATGAGATGTCGTCATGATCATGATTCATGATTCATGATTCATGATTCATG 720  
Db 826 ATTTTATATGAGATGTCGTCATGATCATGATTCATGATTCATGATTCATGATTCATG 885  
QY 721 GTTTTAAATTTGTTGAAGGTGACACACCTTCTCTGCTGAGAGATATCAAAAGACTA 780  
Db 886 GTTTTAAATTTGTTGAAGGTGACACACCTTCTCTGCTGAGAGATATCAAAAGACTA 945  
QY 781 AATGCCATCATGGAAGAGCTGTTGAACAAAGATCCCTTCAATTAAGACATCTGCTATCGAA 840  
Db 946 AATGCCATCATGGAAGAGCTGTTGAACAAAGATCCCTTCAATTAAGACATCTGCTATCGAA 1005  
QY 841 ATTTTAAATTTCCCTTACCTTGTATGAGAGCTGACAGACCTAATGTTGATGATTCAGAA 900  
Db 1006 ATTTTAAAGATCCCTTACCTTGTATGAGAGCTGACAGACCTAATGTTGATGATTCAGAA 1065  
QY 901 ATGACTCTGGAAGACAAAATTTGGATTGTGCAAGGAGGCTGCTCATATAATTAATGCC 960  
Db 1066 ATGACTCTGGAAGACAAAATTTGGATTGTGCAAGGAGGCTGCTCATATAATTAATGCC 1125  
QY 961 ATGCAAAAAGGATCCACCTGAGACTCTGAGGGCACTGTGCAAGATGACAGAAATGAGC 1020  
Db 1126 ATGCAAAAAGGATCCACCTGAGACTCTGAGGGCCCTGTGCAAGATGACAGAAATGAGC 1185  
QY 1021 CCAAGAGAAAGGATGCGCTGAGGAGCTCCAGGCGCTGATGAGAAAGCCAGGAGCTG 1080  
Db 1186 CCAAGAGAAAGGATGAGGCTGAGAAAGCTCCAGGCGCTGATGAGAGCCAGGAGCTG 1245  
QY 1081 AAAAGATTGTGGAAGAAAATATGAAGAAAATAGCAAAAGATGCAAGAAATGAGATCT 1140  
Db 1246 AAAAGATTGTGGAAGAAAATATGAAGAAAATAGCAAAAGATGCAAGAAATGAGATCT 1305  
QY 1141 CGGAATTTTCCAGGAGCTGATGTTGATGTTGATCTCCATGAAACACATTTAAAGGAATG 1200  
Db 1306 CGGAATTTTCCAGGAGCTGATGTTGATGTTGATCTCCATGAAACACATTTAAAGGAATG 1365  
QY 1201 GAAGAAAAGGAGGAGCAACTGAGGAGAGCTTTCTTTTCACTCCCGAGGACGAGATGAA 1260  
Db 1366 GAAGAAAAGGAGGAGCAACTGAGGAGAGCTTTCTTTTCACTCCCGAGGACGAGATGAA 1425  
QY 1261 GAGAGTGGCAAGGAGGAGGAACTGATGAACCAACTTTAGAGAACCTGCTCGAG 1320  
Db 1426 GAGAGTGGCAAGGAGGAGGAACTGATGAACCAACTTTAGAGAACCTGCTCGAG 1485  
QY 1321 TCTCAGCCTATTCTTCCATGAGCTCCAGCACTTGAATCAATTTGAGAGATGCCACA 1380  
Db 1486 CTTGACCTTTTCTTCCATGAGCTCCAGCACTTGAATCAATTTGAGAGATGCCACA 1545  
QY 1381 TCTGACCTTTGATACCATG 1399  
Db 1546 TCTGACCTTTGATATCATG 1564

RESULT 10  
AK123021  
LOCUS  
DEFINITION Homo sapiens cDNA FLJ41025 fis, clone BRACE2008594, weakly similar  
to Human protein serine/threonine kinase stk2 mRNA.  
ACCESSION AK123021  
VERSION AK123021.1 GI:34528469  
KEYWORDS Oligo capping; fis (full insert sequence).  
SOURCE Homo sapiens  
ORGANISM  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
REFERENCE  
1  
Tashiro,H., Yamazaki,M., Watanabe,K., Kumagai,A., Itakura,S.,  
Fukuzumi,Y., Fujimori,Y., Komiya,M., Sugiyama,T., Irie,R.,  
Otsuki,T., Sato,H., Wakamatsu,A., Ishii,S., Yamamoto,J., Isoro,Y.,

Kawai-Hio,Y., Saito,K., Nishikawa,T., Kimura,K., Yamashita,H.,  
Matsuo,K., Nakamura,Y., Sekine,M., Kikuchi,H., Kanda,K.,  
Wagatsuma,M., Murakawa,K., Kanehori,K., Takahashi-Fujii,A.,  
Oshima,A., Sugiyama,A., Kawakami,B., Suzuki,Y., Sugano,S.,  
Nagahara,K., Masuho,Y., Nagai,K. and Isogai,T.  
NEBO human cDNA sequencing project  
Unpublished  
2 (bases 1 to 2189)  
Isogai,T. and Yamamoto,J.  
Direct Submission  
Submitted (15-JUL-2003) Takao Isogai, FLJ Project (HRI Team); 2-6-7  
Kazusa-Kamatari, Kisarazu, Chiba 292-0818, Japan  
(E-mail:genomics@hri.co.jp, Tel:81-438-52-3975, Fax:81-438-52-3986)  
NEBO human cDNA sequencing project supported by Ministry of  
Economy, Trade and Industry of Japan; cDNA full insert sequencing:  
Research Association for Biotechnology (RAB); cDNA library  
construction: Helix Research Institute (HRI) (supported by Japan  
Key Technology Center etc.); 5'- & 3'-end one pass sequencing: RAB,  
HRI, and Biotechnology Center, National Institute of Technology and  
Evaluation; clone selection for full insert sequencing: HRI and  
RAB; annotation: HRI and RAB.  
FEATURES  
source  
1..2189  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="BRACE2008594"  
/tissue\_type="cerebellum"  
/clone\_lib="BRACE2"  
/note="Cloning vector: pME18SFL3"  
ORIGIN  
Query Match 45.0%; Score 651.8; DB 9; Length 2189;  
Best Local Similarity 99.7%; Pred. No. 1.4e-154;  
Matches 653; Conservative 0; Mismatches 2; Indels 0; Gaps 0;  
QY 1 ATGCTGAAATCCAAAGAGGCGAGCTAAGTGTGTGATGGATCAACAGCCATTTCCACTAT 60  
Db 1221 ATGCTGAAATCCAAAGAGGCGAGCTAAGTGTGTGATGGATCAACAGCCATTTCCACTAT 1280  
QY 61 CCAAGACCTTTGATGCAAGAGATAGTGTCTCAACAAAACCTGCGAGTGGAGTTTT 120  
Db 1281 CCAAGACCTTTGATGCAAGAGATAGTGTCTCAACAAAACCTGCGAGTGGAGTTTT 1340  
QY 121 GGAATCTGTATCTGTTTTCAGCAAGAAAGCCAAACGAGGAGAGGAATTAAGGTACTT 180  
Db 1341 GGAATCTGTATCTGTTTTCAGCAAGAAAGCCAAACGAGGAGAGGAATTAAGGTACTT 1400  
QY 181 AAGGAATATCTGTGTGAGAGCTAATCCAAATGAACTGTACAGCCCAATTTGGAGCC 240  
Db 1401 AAGGAATATCTGTGTGAGAGCTAATCCAAATGAACTGTACAGCCCAATTTGGAGCC 1460  
QY 241 CAACTCTCTCCAGCTGGACCCACCCAGCCATTTGTCAAGTTCCATGCAAGTTTGTGGAG 300  
Db 1461 CAACTCTCTCCAGCTGGACCCACCCAGCCATTTGTCAAGTTCCATGCAAGTTTGTGGAG 1520  
QY 301 CAAGATAATTTCTGCATTTATCAGGAGTACTGTGAGGCGCGAGATCTGGACGATAAAT 360  
Db 1521 CAAGATAATTTCTGCATTTATCAGGAGTACTGTGAGGCGCGAGATCTGGACGATAAAT 1580  
QY 361 CAGCAATATAACCAAGCTGGAAAATCTTTCCAGAAAATCAAATATATAGAAATGTTTATC 420  
Db 1581 CAGCAATATAACCAAGCTGGAAAATCTTTCCAGAAAATCAAATATATAGAAATGTTTATC 1640  
QY 421 CAGCTGTCTGAGGAGTTGACTACATGATGAGAGGAGGATATCTCATCGAGACTTAAAG 480  
Db 1641 CAGCTGTCTGAGGAGTTGACTACATGATGAGAGGAGGATATCTCATCGAGACTTAAAG 1700  
QY 481 TCAAGAAATGTAATTTCTGAAAATAATCTCTTAAATTCGAGATTTGGAGTTTCTCGA 540  
Db 1701 TCAAGAAATGTAATTTCTGAAAATAATCTCTTAAATTCGAGATTTGGAGTTTCTCGA 1760  
QY 541 CTTCTAATGGGATCCTGTGACCTGGCCACCAACTTTTAACCTGGAACCTCCCATTTATGAGT 600

```

Db      1761  CTTCTAATGGATCTGTGACCTGGCCACCAACTTTAACTGGAACTCCCATATATAGT 1820
QY      601  CTTGAGGCTCTGAAACACCAAGCTATGACACAAAGTCGGACATCTGTCACGTG 655
Db      1821  CTTGAGGCTCTGAAACACCAAGCTATGACACAAAGTCGGACATCTGTCACGTG 1875

RESULT 11
AC025148/c
LOCUS   AC025148.3 GI:8077076
DEFINITION Homo sapiens chromosome 3 clone RP11-265F19 map 3, WORKING DRAFT
SEQUENCE, 11 unordered pieces.
ACCESSION AC025148
VERSION   AC025148.3
KEYWORDS  HTG; HTGS_PHASE1; HTGS_DRAFT.
SOURCE   Homo sapiens
ORGANISM Homo sapiens
          Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
          Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 154330)
AUTHORS   Birren,B., Linton,L., Nusbaum,C., Lander,E., Abraham,H., Allen,N.,
          Anderson,S., Baldwin,J., Barna,N., Bastien,V., Beda,F.,
          Boguslavskiy,L., Bouckgalter,B., Brown,A., Burkett,G.,
          Campopiano,A., Castle,A., Choepel,Y., Colangelo,M., Collins,S.,
          Collymore,A., Cooke,P., DeArrellano,K., Dewar,K., Diaz,J.S.,
          Dodge,S., Domino,M., Doyle,M., Ferreira,P., FitzHugh,W., Gage,D.,
          Galagan,J., Gardyna,S., Ginde,S., Goyette,M., Graham,L.,
          Grand-Pierre,N., Grant,G., Hagos,B., Heaford,A., Horton,L.,
          Howland,J.C., Iliev,I., Johnson,R., Jones,C., Kann,L., Karatas,A.,
          Klein,J., LaRoque,K., Lamazares,R., Landers,T., Lehoczy,J.,
          Levine,R., Liu,C., Liu,G., Locke,K., Macdonald,P., Marquis,N.,
          McElrim,J., Meneus,L., Mihova,T., Miranda,C., McKernan,K., McPheeters,R.,
          Murphy,T., Naylor,J., Norman,C.H., O'Connor,T., O'Donnell,P.,
          O'Neil,D., Olivier,T.M., Oliver,J., Peterson,K., Pierre,N.,
          Pisani,C., Pollara,V., Raymond,C., Riley,R., Rogov,P., Rothman,D.,
          Roy,A., Santos,R., Schauer,S., Severy,P., Spencer,B.,
          Stange-Thomann,N., Stojanovic,N., Subramanian,A., Talamas,J.,
          Tesfaye,S., Theodore,J., Tirrell,A., Travers,M., Trigilio,J.,
          Vassiliev,H., Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Ye,W.J.,
          Young,G., Zainoun,J., Zimmer,A. and Zody,M.
          Direct Submission
          Submitted (05-MAR-2000) Whitehead Institute/MIT Center for Genome
          Research, 320 Charles Street, Cambridge, MA 02141, USA
REFERENCE 3 (bases 1 to 154330)
AUTHORS   Birren,B., Linton,L., Nusbaum,C., Lander,E., Abraham,H., Allen,N.,
          Anderson,S., Baldwin,J., Barna,N., Bastien,V., Beda,F.,
          Boguslavskiy,L., Bouckgalter,B., Brown,A., Burkett,G.,
          Campopiano,A., Castle,A., Choepel,Y., Colangelo,M., Collins,S.,
          Collymore,A., Cooke,P., DeArrellano,K., Dewar,K., Diaz,J.S.,
          Dodge,S., Domino,M., Doyle,M., Ferreira,P., FitzHugh,W., Gage,D.,
          Galagan,J., Gardyna,S., Ginde,S., Goyette,M., Graham,L.,
          Grand-Pierre,N., Grant,G., Hagos,B., Heaford,A., Horton,L.,
          Howland,J.C., Iliev,I., Johnson,R., Jones,C., Kann,L., Karatas,A.,
          Klein,J., LaRoque,K., Lamazares,R., Landers,T., Lehoczy,J.,
          Levine,R., Liu,C., Liu,G., Locke,K., Macdonald,P., Marquis,N.,
          McElrim,J., Meneus,L., Mihova,T., Miranda,C., McKernan,K., McPheeters,R.,
          Murphy,T., Naylor,J., Norman,C.H., O'Connor,T., O'Donnell,P.,
          O'Neil,D., Olivier,T.M., Oliver,J., Peterson,K., Pierre,N.,
          Pisani,C., Pollara,V., Raymond,C., Riley,R., Rogov,P., Rothman,D.,
          Roy,A., Santos,R., Schauer,S., Severy,P., Spencer,B.,
          Stange-Thomann,N., Stojanovic,N., Subramanian,A., Talamas,J.,
          Tesfaye,S., Theodore,J., Tirrell,A., Travers,M., Trigilio,J.,
          Vassiliev,H., Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Ye,W.J.,
          Young,G., Zainoun,J., Zimmer,A. and Zody,M.
          Direct Submission
          Submitted (24-AUG-2002) Whitehead Institute/MIT Center for Genome
  
```

```

COMMENT
Research, 320 Charles Street, Cambridge, MA 02141, USA
On May 25, 2000 this sequence version replaced gi:7272269.
All repeats were identified using RepeatMasker:
Smit, A.F.A. & Green, P. (1996-1997)
http://ftp.genome.washington.edu/RM/RepeatMasker.html
----- Genome Center
Center: Whitehead Institute/ MIT Center for Genome Research
Center code: WIBR
Web site: http://www-seq.wi.mit.edu
Contact: sequence_submissions@genome.wi.mit.edu
----- Project Information
Center project name: L7980
Center clone name: 265_F_19
----- Summary Statistics
Sequencing vector: M13; M77815; 100% of reads
Chemistry: Dye-terminator Big Dye; 100% of reads
Assembly program: Phrap; version 0.960731
Consensus quality: 149112 bases at least Q40
Consensus quality: 151109 bases at least Q30
Consensus quality: 152128 bases at least Q20
Insert size: 150000; agarose-fp
Quality coverage: 5.6 in Q20 bases; agarose-fp
Quality coverage: 5.5 in Q20 bases; sum-of-contigs
-----
* NOTE: This is a 'working draft' sequence. It currently
* consists of 11 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.
*
* 1 1829: contig of 1829 bp in length
* 1830: gap of 100 bp
* 1930: contig of 3622 bp in length
* 5551: gap of 100 bp
* 5552: contig of 6148 bp in length
* 11799: gap of 100 bp
* 11800: contig of 8827 bp in length
* 20726: gap of 100 bp
* 20827: contig of 12335 bp in length
* 33161: gap of 100 bp
* 33162: contig of 13359 bp in length
* 33262: gap of 100 bp
* 46721: contig of 16615 bp in length
* 46721: gap of 100 bp
* 63336: gap of 100 bp
* 63436: gap of 100 bp
* 81257: gap of 100 bp
* 81357: contig of 19169 bp in length
* 100526: gap of 100 bp
* 100526: contig of 22201 bp in length
* 122826: gap of 100 bp
* 122827: contig of 31404 bp in length.
* 122927: Location/Qualifiers
* 1.154330
* /organism="Homo sapiens"
* /mol_type="genomic DNA"
* /db_xref="taxon:9606"
* /chromosome="3"
* /map="3"
* /clone="RP11-265F19"
* /clone_lib="RPC1-11 Human Male BAC"
* 1.1829
* /note="assembly_fragment"
* 1930..5551
* /note="assembly_fragment"
* 5552..11799
* /note="assembly_fragment"
* 11900..20726
* /note="assembly_fragment"
* 20827..33161
* /note="assembly_fragment"
  
```

```
misc_feature 33262..46620
/note="assembly_fragment"
misc_feature 46721..63335
/note="assembly_fragment"
misc_feature 63436..81256
/note="assembly_fragment"
clone_end:SP6
vector_side:right"
misc_feature 81357..100525
/note="assembly_fragment"
misc_feature 100626..122826
/note="assembly_fragment"
clone_end:TV
vector_side:left"
misc_feature 122927..154330
/note="assembly_fragment"

ORIGIN
Query Match 11.9%; Score 172.2; DB 2; Length 154330;
Best Local Similarity 98.3%; Pred. No. 1-5e-32;
Matches 174; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 ATGCTGAATTCACAGGAGCAGTAAGTGTGTGAGTGATCAACAGCCATTCCACTTAT 60
Db 137522 ATGCTGAATTCACAGGAGCAGTAAGTGTGTGAGTGATCAACAGCCATTCCACTTAT 137463

QY 61 CCAAGACCTTGATTCACAGAGATACGCTTCAACAAAACCTGGCAGTGGAAAGTTT 120
Db 137462 CCAAGACCTTGATTCACAGAGATACGCTTCAACAAAACCTGGCAGTGGAAAGTTT 137403

QY 121 GGAACCTGTCTATCTGTTTCAGACAGAAAGCCAAACGAGGAGAGGAATTAAGGTA 177
Db 137402 GGAACCTGTCTATCTGTTTCAGACAGAAAGCCAAACGAGGAGAGGAATTAAGGTA 137346

RESULT 12
AC055733/c
LOCUS Homo sapiens 3 BAC RP11-39E3 (Roswell Park Cancer Institute Human
DEFINITION BAC Library) complete sequence.
ACCESSION AC055733
VERSION AC055733.16 GI:18463999
KEYWORDS Htg.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 193425)
Muzny,D.M., Adams,C., Adio-Oduola,B., Ali-oman,F.R., Allen,C.,
Alsbrooks,S.L., Amaralunge,H.C., Are,J.R., Ayete,M., Banks,T.,
Barbaria,J., Benton,J., Binage,K., Blankenburg,K., Bonnin,D.,
Bouck,J., Bowie,S., Brieva,M., Brown,E., Brown,M., Bryant,N.P.,
Buhay,C., Burch,P., Burkett,C., Burrell,K.L., Byrd,N.C.,
Carson,T.F., Carter,M., Cavazos,S.R., Chacko,J., Chavez,D.,
Chen,G., Chen,R., Chen,Z., Chowdhry,I., Christopoulos,C.,
Cleveland,C.D., Cox,C., Coyle,M.D., Dathorne,S.R., David,R.,
Davila,M.L., Davis,C., Davy-Carroll,L., Dederich,D.A.,
Delaney,K.R., Delgado,O., Penn,A.L., Ding,Y., Dinh,H.H.,
Douthwaite,K.J., Draper,H., Dugan-Rocha,S., Durbin,K.J.,
Earnhart,C., Edgar,D., Edwards,C.C., Elhaj,C., Escotto,M.,
Falls,T., Ferraguto,D., Flagg,N., Ford,J., Foster,P., Frantz,P.,
Gabisi,A., Gao,J., Garcia,A., Garner,T., Garza,N., Gill,R.,
Gorelli,J.H., Guevara,W., Gunaratne,P., Hale,S., Hamilton,K.,
Harris,C., Harris,K., Hart,M., Haviak,P., Hawes,A., He,X.,
Hernandez,J., Hernandez,O., Hodgson,A., Hogue,M., Holloway,C.,
Hollins,B., Homs,F., Howard,S., Huber,J., Hulyk,S., Hume,J.,
Jackson,L.E., Jacobson,B., Jia,Y., Johnson,R., Jolivet,S.,
Joudah,S., Karlsson,E., Kelly,S., Khan,U., King,L., Korvah,J.,
Kovar,C., Kratovic,J., Kureshi,A., Landry,N., Leal,B., Lewis,L.C.,
Lewis,L., Li,J., Li,Z., Lichtarge,O., Lieu,C., Liu,J., Liu,W.,
Louiseged,H., Lozado,R.J., Lu,X., Lucier,A., Lucier,R., Luna,R.,
Ma,J., Maheshwari,M., Mapua,P., Martin,R., Martindale,A.,
Martinez,E., Massey,E., Mawhinney,E., McLeod,M.P., Meador,M.,
```

```
Mei,G., Metzker,M., Miner,G., Miner,Z., Mitchell,T., Mohabbat,K.,
Moore,S., Morgan,M., Moorish,T., Morris,S., Moser,M., Neal,D.,
Nelson,D., Newton,J., Newton,N., Nguyen,A., Nguyen,N., Nguyen,N.,
Nickerson,E., Nwokenkwo,S., Ogih,M., Okwuonu,G., Oragunye,N.,
Oviedo,R., Pace,A., Payton,B., Peery,J., Perez,L., Peters,M.,
Pickens,A., Primus,E., Pu,L.L., Quiles,M., Ren,Y., Rives,M.,
Rojas,A., Rojebokan,I., Rolfe,M., Ruiz,S., Savery,G., Scherer,S.,
Scott,G., Shen,H., Shoshitari,N., Sisson,I., Sodargren,E.,
Sonaik,T., Sparks,A., Stanley,H., Stone,H., Sutton,A., Svatek,A.,
Tabor,P., Tamerisa,A., Tamerisa,K., Tang,H., Tansey,J., Taylor,C.,
Taylor,T., Telford,B., Thomas,N., Thomas,S., Usmani,K., Vasquez,L.,
Vera,V., Villalon,D., Vinson,R., Wall,R., Wang,S., Ward-Moore,S.,
Warren,R., Washington,C., Watlington,S., Williams,G.,
Williamson,A., Wlezyk,R., Wooden,S., Worley,K., Wu,C., Wu,Y.,
Wu,Y.F., Zhou,J., Zorrilla,S., Zorrilla,S., Zorrilla,S., Zorrilla,S.,
Gibbs,R.

Direct Submission
Unpublished
2 (bases 1 to 193425)
Worley,K.C.
Direct Submission
Submitted (18-APR-2000) Human Genome Sequencing Center, Department
of Molecular and Human Genetics, Baylor College of Medicine, One
Baylor Plaza, Houston, TX 77030, USA
3 (bases 1 to 193425)
Worley,K.C.
Direct Submission
Submitted (01-FEB-2002) Human Genome Sequencing Center, Department
of Molecular and Human Genetics, Baylor College of Medicine, One
Baylor Plaza, Houston, TX 77030, USA
4 (bases 1 to 193425)
Worley,K.C.
Direct Submission
Submitted (28-MAR-2002) Human Genome Sequencing Center, Department
of Molecular and Human Genetics, Baylor College of Medicine, One
Baylor Plaza, Houston, TX 77030, USA
5 (bases 1 to 193425)
Worley,K.C.
Direct Submission
Submitted (30-JUL-2002) Human Genome Sequencing Center, Department
of Molecular and Human Genetics, Baylor College of Medicine, One
Baylor Plaza, Houston, TX 77030, USA
On Feb 1, 2002 this sequence version replaced gi:17432818.
INFORMATION: http://www.hgsc.bcm.tmc.edu/ or email
gc-help@bcm.tmc.edu

CLONE LENGTH: This sequence does not necessarily represent the
entire insert of this clone. Overlapping regions of clones are only
sequenced and submitted once, so the sequence for the remainder of
the insert may be found in the record for the adjacent clones.
Overlapping clones are noted at the beginning and end of the
Features listing.

ANNOTATION OF FEATURES:
STSS are identified using ePCR (Genome Res. 7:541-550) searches
of a local database that includes entries from dbSTS, GDB, and
local mapping efforts.
Repeats are identified using RepeatMasker (A. Smit and P. Green,
unpublished.) for Human and Mouse sequences.
Genes and Region of sequence similarity are identified by BLAST
(Nuc. Acids Res. 25:3389-3402) similarity (expect < 1e-34) to the
EST and cDNA sequences. Genes demonstrate at least two exons
flanked by consensus splice sites that maintained sequence
continuity across the splice junctions. Sequences that are not
identical matches are annotated as similar.

SEQUENCING READ COVERAGE: Sequencing is completed to a minimum
standard of double strand coverage with a minimum of 2 clones and 2
reads with no ambiguities or 2 chemistries with a minimum of 2
clones and 3 reads, with no ambiguities. If the sequence quality for
a region does not meet this standard, it will be indicated in the
annotation as Low Coverage.
```



Martinez,E., Massey,E., Mawhiney,E., McLeod,M.P., Meador,M., Mei,G., Metzker,M., Miner,G., Miner,Z., Mitchell,T., Mohabbat,K., Moore,S., Morgan,M., Moorish,T., Morris,S., Moser,M., Neal,D., Nelson,D., Newton,J., Newton,N., Nguyen,A., Nguyen,N., Nguyen,N., Nickerson,E., Nwokweto,S., Oguh,M., Okwono,G., Oragunye,N., Oviedo,R., Pace,A., Payton,B., Peery,J., Perez,L., Peters,L., Pickens,R., Primus,E., Pu,L.L., Quiles,M., Ren,Y., Rives,M., Rojias,A., Rojubokan,I., Rolfe,M., Ruiz,S., Savery,G., Scherer,S., Scott,G., Shen,H., Shoohtari,N., Sisson,I., Sodergren,E., Sotak,A., Sparks,A., Stanley,H., Stone,H., Sutton,A., Svatek,A., Taber,P., Tamerisa,A., Tamerisa,K., Tang,H., Tansey,J., Taylor,C., Taylor,T., Telford,B., Thomas,N., Thomas,S., Usmani,K., Vasquez,L., Vera,V., Villalon,D., Vinson,R., Wang,Q., Wang,S., Ward-Moore,S., Warren,R., Washington,C., Watlington,S., Williams,G., Williamson,A., Wleczky,R., Wooden,S., Worley,K., Wu,C., Wu,Y., Wu,Y.F., Zhou,J., Zorrilla,S., Naylor,S.L., Weinstein,G. and Gibbs,R.

Direct Submission  
Unpublished  
2 (bases 1 to 60398)  
Worley,K.C.

Direct Submission  
Submitted (28-MAR-2002) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA  
3 (bases 1 to 60398)  
Worley,K.C.

Direct Submission  
Submitted (30-MAR-2002) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA  
4 (bases 1 to 60398)  
Worley,K.C.

Direct Submission  
Submitted (28-JUN-2002) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA  
INFORMATION: <http://www.hgsc.bcm.tmc.edu/> or email [gc-help@bcm.tmc.edu](mailto:gc-help@bcm.tmc.edu)

CLONE LENGTH: This sequence does not necessarily represent the entire insert of this clone. Overlapping regions of clones are only sequenced and submitted once, so the sequence for the remainder of the insert may be found in the record for the adjacent clones. Overlapping clones are noted at the beginning and end of the Features listing.

#### ANNOTATION OF FEATURES:

STSs are identified using ePCR (Genome Res. 7:541-550) searches of a local database that includes entries from dbSTS, GDS, and local mapping efforts.

Repeats are identified using RepeatMasker (A. Smit and P. Green, unpublished.) for Human and Mouse sequences.

(Genes and Region of sequence similarity are identified by BLAST (Nuc. Acids Res. 25:3389-3402) similarity (expect < 1e-34) to the EST and cDNA sequences. Genes demonstrate at least two exons flanked by consensus splice sites that maintained sequence continuity across the splice junctions. Sequences that are not identical matches are annotated as similar.

SEQUENCING READ COVERAGE: Sequencing is completed to a minimum standard of double strand coverage with a minimum of 2 clones and 2 reads with no ambiguities or 2 chemistries with a minimum of 2 clones and 3 reads with no ambiguities. If the sequence quality for a region does not meet this standard, it will be indicated in the annotation as Low Coverage.

QUALITY OF INDIVIDUAL BASES: This sequence meets stringent quality standards - estimated error rate less than 1 per 10,000 bases. Reports of lowest quality individual bases and measures of base quality are listed below. Description of the metrics can be found at URL:

<http://gc.bcm.tmc.edu:8088/quality.info/genbank.annotation.html>.

FEATURES	QUALSTAT-REPORT.	Location/Qualifiers
Source	1. 60398	/organism="Homo sapiens"
		/mol_type="Genomic DNA"
		/db_xref="taxon:9606"
		/chromosome="3"
		/clones="RP11-265F19"
misc_feature	complement(1..2005)	/note="Overlaps bases 1..2005 of clone AC121332"
	/function="clone overlap"	
repeat_region	13..1803	/rpt_family="L1P4"
repeat_region	complement(1804..2070)	/rpt_family="AluSg"
repeat_region	2071..3576	/rpt_family="L1P4"
repeat_region	3577..3807	/rpt_family="AluSg/x"
repeat_region	3808..3996	/rpt_family="L1P4"
repeat_region	complement(3976..4781)	/rpt_family="L1P4"
STS	4444..4621	/standard_name="61535"
repeat_region	4783..5641	/rpt_family="L1P4"
repeat_region	5644..5680	/rpt_family="L1P4"
repeat_region	5681..6309	/rpt_family="L1P4"
repeat_region	6433..7738	/rpt_family="L1P4"
repeat_region	8070..8214	/rpt_family="MER57-internal"
repeat_region	8241..8544	/rpt_family="MER57-internal"
repeat_region	8680..8954	/rpt_family="MER57-internal"
repeat_region	9458..9582	/rpt_family="AluSg"
repeat_region	9577..10252	/rpt_family="U6"
repeat_region	10284..10645	/rpt_family="MER57-internal"
repeat_region	10999..11020	/rpt_family="MER57B"
repeat_region	11322..11627	/rpt_family="AT-rich"
repeat_region	12348..12369	/rpt_family="AluJo"
STS	12449..12646	/rpt_family="CAAAA"
repeat_region	12851..12957	/standard_name="173989"
repeat_region	complement(13041..13219)	/rpt_family="AluSx"
repeat_region	14844..15152	/rpt_family="MER5A"
repeat_region	15679..15779	/rpt_family="AluSc"
repeat_region	15782..16643	/rpt_family="MIR"
repeat_region	16688..16826	/rpt_family="L1MC4"
repeat_region	17461..17482	/rpt_family="L2"
repeat_region	17487..17655	/rpt_family="AT-rich"
repeat_region	complement(18331..18445)	/rpt_family="MER5B"
repeat_region	complement(19199..19389)	/rpt_family="L2"



```

/rpt_family="MER5A"
19396..19676
/rpt_family="LMC4"
complement(19733..19937)
/rpt_family="MER3"
complement(21247..21557)
/rpt_family="AluJo"
22046..22393
/standard_name="173331"
22175..22391
/standard_name="61389"
22252..22391
/standard_name="59650"
22284..22315
/rpt_family="(CA)n"
22316..22359
/rpt_family="(CA)n"
23450..23482
/rpt_family="AT_rich"

Query Match 11.5%; Score 167.2; DB 9; Length 60398;
Best Local Similarity 98.3%; Pred. No. 2.6e-31;
Matches 169; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 171 AAAGGTACTTAAGGAATATCTTGGAGAACTTAATCAATGAACTGTACAGGCCAA 230
Db 21962 AAAGGTACTTAAGGAATATCTTGGAGAACTTAATCAATGAACTGTACAGGCCAA 21903

Qy 231 TTGTGAAGCCAACTCTCTCCAAAGTGGACCAACCAGCCATGTCAAGTTCATGCAAG 290
Db 21902 TTGTGAAGCCAACTCTCTCCAAAGTGGACCAACCAGCCATGTCAAGTTCATGCAAG 21843

Qy 291 TTGTGTGGACGACAGATAATTTCTGCATTATCAAGGAGTACTGTGAGGGCCGA 342
Db 21842 TTGTGTGGACGACAGATAATTTCTGCATTATCAAGGAGTACTGTGAGGTGAGA 21791

RESULT 14
AC083908 111568 bp DNA linear PRI 28-JUN-2002
LOCUS Homo sapiens 3 BAC RP11-431K13 (Roswell Park Cancer Institute Human
DEFINITION BAC library) complete sequence.
ACCESSION AC083908
VERSION AC083908.19 GI:21263155
KEYWORDS HTG.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 111568)
Muzny,D.M., Adams,C., Adio-Oduola,B., Ali-osman,F.R., Allen,C.,
Alsbrooks,S.L., Amarantunge,H.C., Are,J.R., Ayale,M., Banks,T.,
Barbata,J., Benton,J., Bimaga,K., Blankenburg,K., Bonnin,D.,
Bouck,J., Bowie,S., Brieva,M., Brown,E., Brown,M., Bryant,N.P.,
Buhay,C., Burch,P., Burkett,C., Burrell,K.L., Byrd,N.C.,
Carroll,T.F., Carter,M., Cavazos,S.R., Chacko,J., Chavez,D.,
Chen,R., Chen,R., Chen,Z., Chowdhry,I., Christopoulos,C.,
Cleveland,C.D., Cox,C., Coyle,M.D., Dathorne,S.R., David,R.,
Davila,M.L., Davis,C., Davy-Carroll,L., Dederich,D.A.,
Delaney,K.R., Delgado,O., Dunn,A.L., Ding,Y., Dinh,H.H.,
Douthwaite,K.J., Draper,H., Dugan-Rocha,S., Durbin,K.J.,
Earnhart,C., Edgar,D., Edwards,C.C., Elhaj,C., Escotto,M.,
Falls,T., Ferraguto,D., Flagg,N., Ford,J., Foster,P., Frantz,P.,
Gabisi,A., Gao,J., Garcia,A., Garner,T., Garza,N., Gill,R.,
Gorrell,J.H., Guevara,W., Gunaratne,P., Hale,S., Hamilton,K.,
Harris,C., Harris,K., Hart,M., Haylak,P., Hawes,A., He,X.,
Hernandez,J., Hernandez,O., Hodgson,A., Hughes,M., Holloway,C.,
Hollins,B., Honsi,F., Howard,S., Huber,J., Hulyk,S., Hume,J.,
Jackson,L.E., Jacobson,B., Jia,Y., Johnson,R., Jolivet,S.,
Joudan,S., Karlsson,E., Kelly,S., Khan,U., King,L., Korvah,J.,
Kovar,C., Kratovic,J., Kureshi,A., Landry,N., Leal,B., Lewis,L.C.,
Lewis,L., Li,J., Li,Z., Lichtarge,O., Lieu,C., Liu,J., Liu,W.,
Loulseged,H., Lozado,R.J., Lu,X., Lucier,A., Lucier,R., Luna,R.,

```

```

Ma,J., Maheshwari,M., Mapua,P., Martin,R., Martindale,A.,
Martinez,E., Massey,E., Mawhney,E., McLeod,M.P., Meador,M.,
Mei,G., Metzger,M., Miner,G., Miner,Z., Mitchell,T., Monabbat,K.,
Moore,S., Morgan,M., Moorish,T., Morris,S., Moser,M., Neal,D.,
Nelson,D., Newton,J., Newton,N., Nguyen,A., Nguyen,N., Nguyen,N.,
Nickerson,E., Nwokenkwo,S., Oguh,M., Okwuonu,G., Oragunye,N.,
Oviedo,R., Pace,A., Payton,B., Peery,J., Perez,L., Peters,L.,
Pickens,R., Primus,E., Pu,L.L., Quiles,M., Ren,Y., Rives,M.,
Rojas,A., Rojubokan,I., Rolfe,M., Ruiz,S., Savery,G., Scherer,S.,
Scott,G., Shen,H., Shooshtari,N., Sisson,I., Sodergren,E.,
Sonaika,T., Sparks,A., Stanley,H., Stone,H., Sutton,A., Svatek,A.,
Tabor,P., Tamerisa,A., Tamerisa,K., Tang,H., Tansey,J., Taylor,C.,
Taylor,T., Telford,B., Thomas,N., Thomas,S., Usmani,K., Vasquez,L.,
Vera,V., Villalon,D., Vinson,R., Wang,Q., Wang,S., Ward-Moore,S.,
Warren,R., Washington,C., Watlington,S., Williams,G.,
Williamson,A., Wleczyk,R., Wooden,S., Worley,K., Wu,C., Wu,Y.,
Wu,Y.F., Zhou,J., Zorrilla,S., Naylor,S.L., Weinstein,G. and
Gibbs,R.
Direct Submission
Unpublished
2 (bases 1 to 111568)
Worley,K.C.
Direct Submission
Submitted (06-OCT-2000) Human Genome Sequencing Center, Department
of Molecular and Human Genetics, Baylor College of Medicine, One
Baylor Plaza, Houston, TX 77030, USA
3 (bases 1 to 111568)
Worley,K.C.
Direct Submission
Submitted (21-MAY-2002) Human Genome Sequencing Center, Department
of Molecular and Human Genetics, Baylor College of Medicine, One
Baylor Plaza, Houston, TX 77030, USA
4 (bases 1 to 111568)
Worley,K.C.
Direct Submission
Submitted (30-MAY-2002) Human Genome Sequencing Center, Department
of Molecular and Human Genetics, Baylor College of Medicine, One
Baylor Plaza, Houston, TX 77030, USA
5 (bases 1 to 111568)
Worley,K.C.
Direct Submission
Submitted (28-JUN-2002) Human Genome Sequencing Center, Department
of Molecular and Human Genetics, Baylor College of Medicine, One
Baylor Plaza, Houston, TX 77030, USA
On May 30, 2002 this sequence version replaced gi:20976448.
INFORMATION: http://www.hgsc.bcm.tmc.edu/ or email
gc-help@bcm.tmc.edu

```

CLONE LENGTH: This sequence does not necessarily represent the entire insert of this clone. Overlapping regions of clones are only sequenced and submitted once, so the sequence for the remainder of the insert may be found in the record for the adjacent clones. Overlapping clones are noted at the beginning and end of the Features listing.

#### ANNOTATION OF FEATURES:

STPs are identified using ePCR (Genome Res. 7:541-550) searches of a local database that includes entries from dbSTS, GDB, and local mapping efforts.

Repeats are identified using RepeatMasker (A. Smit and P. Green, unpublished) for Human and Mouse sequences.

Genes and Region of sequence similarity are identified by BLAST (Nuc. Acids Res 25:3389-3402) similarity (expect < 1e-34) to the EST and cDNA sequences. Genes demonstrate at least two exons flanked by consensus splice sites that maintained sequence continuity across the splice junctions. Sequences that are not identical matches are annotated as similar.

SEQUENCING READ COVERAGE: Sequencing is completed to a minimum standard of double strand coverage with a minimum of 2 clones and 2 reads with no ambiguities or 2 chemistries with a minimum of 2 clones and 3 reads with no ambiguities. If the sequence quality for a region does not meet this standard, it will be indicated in the





Db 104773 TCCTGAGAGATATCCAAAGAAGAACTAAATGCCATCATGGAAAG 104731

Search completed: October 2, 2004, 23:48:58  
Job time : 5798 secs

----- Project Information  
Center project name: 888  
Center clone name: RP11-290D10  
----- Summary Statistics  
Sequencing Vector: M13mp18; X02513; 100% of reads  
Chemistry: Dye-terminator Big Dye; 100% of reads  
Sequencing Vector: plasmid; plasmid\_accession; 0% of reads  
Chemistry: Dye-terminator Big Dye; 100% of reads  
Assembly program: Phrap; version 0.990319  
Consensus quality: 171536 bases at least Q40  
Consensus quality: 173017 bases at least Q30  
Consensus quality: 173757 bases at least Q20  
Insert size: 16599; agarose-fp  
Insert size: 17542; sum-of-contigs

Quality coverage: 11.9x in Q20 bases; agarose-fp  
Quality coverage: 11.5x in Q20 bases; sum-of-contigs.  
\* NOTE: This is a 'working draft' sequence. It currently  
\* consists of 6 contigs. The true order of the pieces  
\* is not known and their order in this sequence record is  
\* arbitrary. Gaps between the contigs are represented as  
\* runs of N, but the exact sizes of the gaps are unknown.  
\* This record will be updated with the finished sequence.  
\* as soon as it is available and the accession number will  
\* be preserved.

1 1839: contig of 1839 bp in length  
\* 1840: gap of unknown length  
\* 1840 3792: contig of 1853 bp in length  
\* 3793 3892: gap of unknown length  
\* 3893 6447: contig of 2555 bp in length  
\* 6448 6547: gap of unknown length  
\* 6548 8332: contig of 1785 bp in length  
\* 8333 8432: gap of unknown length  
\* 8433 25713: contig of 17281 bp in length  
\* 25714 25813: gap of unknown length  
\* 25814 175923: contig of 150110 bp in length.

## FEATURES

Location/Qualifiers  
1. .175923  
source  
/organism="Homo sapiens"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:9606"  
/chromosome="3"  
/clone="RP11-290D10"  
/clone\_lib="RPCI human BAC library 11"  
1. .1839  
/note="assembly\_name:Contig36"  
misc\_feature 1840. .3792  
/note="assembly\_name:Contig41"  
misc\_feature 3893. .6447  
/note="assembly\_name:Contig44"  
misc\_feature 6548. .8332  
/note="assembly\_name:Contig46"  
misc\_feature 8433. .25713  
/note="assembly\_name:Contig47"  
clone\_end:77"  
misc\_feature 25814. .175923  
/note="assembly\_name:Contig48"  
clone\_end:SP6"

## ORIGIN

Query Match 10.6%; Score 153.4; DB 2; Length 175923;  
Best Local Similarity 96.3%; Pred. No. 9.3e-28;  
Matches 157; Conservative 0; Mismatches 6; Indels 0; Gaps 0;  
QY 635 AGTCGACATCTCGTCACATGCGATGATTTGATGAGATGCTGCGATCATGCAAT 694  
Db 104893 ACITTTGCTTCAGGTCACATGCGATGATTTGATGAGATGCTGCGATCATGCAAT 104834  
QY TCCTGGCTCCATTTCTTATCCATTTGTTTAAATTTGTTGAAGTGACACACCTTCTC 754  
Db 104833 TCCTGGCTCCATTTCTTATCCATTTGTTTAAATTTGTTGAAGTGACACACCTTCTC 104774  
QY 755 TCCTGAGAGATATCCAAAGAAGAACTAAATGCCATCATGGAAAG 797

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: October 2, 2004, 17:05:50 ; Search time 4039 seconds

(without alignments)  
10713.122 Million cell updates/sec

Title: US-10-803-278-3

Perfect score: 1449

Sequence: 1 atgctgaattccaagagc.....actggaaaatgaaaaataa 1449

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 27513289 seqs, 14931090276 residues

Total number of hits satisfying chosen parameters: 55026578

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

EST:\*

1: em\_estba:\*

2: em\_esthum:\*

3: em\_estin:\*

4: em\_estmu:\*

5: em\_estov:\*

6: em\_estpl:\*

7: em\_estro:\*

8: em\_htc:\*

9: gb\_est1:\*

10: gb\_est2:\*

11: gb\_est3:\*

12: gb\_est4:\*

13: gb\_est5:\*

14: gb\_est5:\*

15: em\_estfun:\*

16: em\_estom:\*

17: em\_gss\_hum:\*

18: em\_gss\_inv:\*

19: em\_gss\_pln:\*

20: em\_gss\_vrt:\*

21: em\_gss\_fun:\*

22: em\_gss\_mam:\*

23: em\_gss\_mus:\*

24: em\_gss\_pro:\*

25: em\_gss\_rod:\*

26: em\_gss\_pbg:\*

27: em\_gss\_vrl:\*

28: gb\_gss1:\*

29: gb\_gss2:\*

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1310	90.4	2880	11	BC026195
2	854	58.9	2584	11	AK030042
3	852.4	58.8	1784	11	AK043543
4	841.4	58.1	1201	13	BX365473

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

5	809.8	55.9	1221	11	AK054237
6	714.4	49.3	937	13	BQ709169
7	692.4	47.8	932	13	EX328072
8	618.8	42.7	2239	11	AK030186
9	614	42.4	674	11	BC009414
10	570.2	39.4	671	10	BE388672
11	564.6	39.0	659	12	BE752438
12	510.6	35.2	767	12	BI544364
13	509.6	35.2	867	13	BU930538
14	504	34.8	845	14	CD106568
15	465.2	32.1	635	10	BE552498
16	462.2	31.9	849	13	EX428086
17	388	26.8	878	14	CD557726
18	356.2	24.6	1148	12	EM563364
19	354.4	24.5	609	10	BB666177
20	337.2	23.3	584	14	CD689119
21	336.4	23.2	1126	12	EM563363
22	325.2	22.4	713	12	EG744158
23	315.8	21.8	480	9	AI384010
24	313.2	21.6	628	10	BB616480
25	296.8	20.5	374	9	AI288640
26	273.2	18.9	776	12	BG718517
27	272.4	18.8	620	10	BB617055
28	240	16.6	616	10	BE262600
29	215.4	14.9	675	14	CD251322
30	210.8	14.5	612	13	BU674219
31	199	13.7	525	9	AA010796
32	178.8	12.3	471	12	BI032921
33	176.8	12.2	183	10	BF934127
34	171.6	11.8	609	13	BM197769
35	161.8	11.2	748	29	PT010C03R
36	154.4	10.7	500	12	BF190100
37	154.2	10.6	592	13	BW240639
38	149.8	10.3	586	29	CE484348
39	149.6	10.3	1522	11	BC015515
40	149.6	10.3	2206	29	AY403632
41	147.8	10.2	702	12	BI754945
42	147.8	10.2	913	13	BQ432111
43	147.8	10.2	1201	13	EX364188
44	146.4	10.1	2136	29	AY403634
45	146.4	10.1	3193	11	AK078809

#### ALIGNMENTS

RESULT 1  
BC026195  
LOCUS BC026195 2880 bp mRNA linear HTC 19-NOV-2003  
DEFINITION Homo sapiens NIMA (never in mitosis gene a)-related kinase 11,  
mRNA (CDNA clone IMAGE:4828434), containing frame-shift errors.  
ACCESSION BC026195  
VERSION BC026195.1 GI:20072185  
KEYWORDS HTC  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
REFERENCE 1 (bases 1 to 2880)  
AUTHORS Strausberg, R.L., Feingold, E.A., Grouse, L.H., Derge, J.G., Klausner, R.D., Collins, F.S., Wagner, L., Shenmen, C.M., Schuler, G.D., Altschul, S.F., Zeeberg, B., Buetow, K.H., Schaefer, C.F., Bhat, N.K., Hopkins, R.F., Jordan, H., Moore, T., Max, S.I., Wang, J., Hsieh, F., Diatchenko, L., Marusina, K., Farmer, A.A., Rubin, G.M., Hong, L., Stapleton, M., Soares, X.B., Bonaldo, M.F., Casavant, T.L., Scheetz, T.E., Brownstein, M.J., Usdin, T.B., Toshiyuki, S., Carninci, P., Prange, C., Raha, S.S., Loquellano, N.A., Peters, G.J., Abramson, R.D., Mullahy, S.J., Bosak, S.A., McEwan, P.J., McKernan, K.J., Malek, J.A., Gunatratne, P.H., Richards, S., Worley, K.C., Hale, S., Garcia, A.M., Gay, L.J., Hulyk, S.W., Villalón, D.K., Muzny, D.M., Sodergren, E.J., Lu, X., Gibbs, R.A., Fahey, J., Helton, E., Kettman, M., Madan, A., Rodriguez, S., Sanchez, A., Whitting, M., Madan, A., Young, A.C., Shevchenko, Y.,



Submitted (16-JUL-2001) Yoshihide Hayashizaki, The Institute of Physical and Chemical Research (RIKEN), Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center (GSC), RIKEN Yokohama Institute; 1-7-22 Suehiro-cho, Tsurumi-Ku, Yokohama, Kanagawa 230-0045, Japan (E-mail: genome-res@gsc.riken.go.jp, URL: <http://genome.gsc.riken.go.jp/>, Tel: 81-45-503-9222, Fax: 81-45-503-9216)

cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in RIKEN. Division of Experimental Animal Research in Riken contributed to prepare mouse tissues.

Please visit our web site for further details.  
 URL: <http://genome.gsc.riken.go.jp/>  
 URL: <http://fantom.gsc.riken.go.jp/>  
 Location/Qualifiers

```

1. 2584
    /organism="Mus musculus"
    /mol_type="mRNA"
    /strain="C57BL/6J"
    /db_xref="FANTOM_DB:4932416N14"
    /db_xref="MGI:2392631"
    /db_xref="taxon:10090"
    /clone="4932416N14"
    /sex="male"
    /tissue type="testis"
    /clone_lib="RIKEN full-length enriched mouse cDNA library"
    /dev_stage="adult"
    191..2077
    /notes="unnamed protein product; putative
    similar to MRNA, SIMILAR TO HUMAN HYPOTHETICAL PROTEIN
    FLJ23495, COMPLETE CDS [Macaca fascicularis]
    (SPTB|BAB83539, evidence: FASTA, 72.4%ID, 1008length,
    match=1806)"
    /codon_start=1
    /protein_id="BAC26756.1"
    /db_xref="GI:26326025"
    /translation="MLKFOETAKCVGRRTTVPYPTALTARYVQLQKLGSGFGTV
    YLVSDIKKXPGBELKYLKISVGEKLPNETVOANVEAQLSLRHHPDAVRFHAFSMEQ
    ENFICITTECEGRDLDTYQYKEAGKQVAENQIVWFQILLGVDDVMEHRLILHDL
    KSKNIFLKNLXLKIGDFGVSRLLMSGSELATILTGPHYMSPEALKHQGDYDAKSDIWS
    LACIAEMCLDFHAFAGSFLSVNLNIVEGKTSPLDPRPRELNTIMERLNKSPSLR
    PSAADILKPAWEEQQLQMLKQPMETLKNVSOQKEAAHTINAVQKLLHLQTLQAL
    SQTQTPPRMWRKLQADRARLRKIAENYKENDRKQALSRNVGSHVAVL
    HELDTRLSLPDLPSQLDLDELEPSLEDITVLGLHVEIIPDPLVAEQYVSDVDFK
    CEDSSEOEEMIFPSAGDQTEKESPSVVTNQDSDTAAALVGLCHLVLYGTVSLDTK
    TLTNAVDSRPGVWNSAVATKMKEMKESAVQKLGMETFPEVVDYLYKRRARRQNE
    AINWEHLETVPRASDCFVDQLLIFEEILLTWEGKEPFLQNLPCBAQKPKPVKGTGH
    CNP"
```

ORIGIN

Query Match	58.9%;	Score 854;	DB 11;	Length 2584;
Best Local Similarity	83.2%;	Pred. No. 4.8e-193;		
Matches	984;	Conservative	0;	Mismatches 195; Indels 3; Gaps 1;

Qy	1	ATGCTGAATTC	CAAGGCGCAGCTAAGTGTGTG---AGTGGATCAACAGCCATTTCAACT	57
Db	191	ATGCTGAATTC	CAAGAGACTCGGAAGTGTGTGGGAGAGAACCCACAGCTCATCCCATG	250
Qy	58	TATCCAAAGACCTT	GTATTCGAAGAAGATACGTCGTTTCAACAAAAATCTCGCAGTGGCAAGT	117
Db	251	TACCGACCGCTT	GTATTCGAAGAAGATATGTCTTCACAGAAATCTGGCAGTGGAAAT	310
Qy	118	TTTGGAACTGTCT	ATCTGTTTTCAGACAAGAAAGCCAAACGAGGAGAGGAATTTAAAGGTA	177
Db	311	TTTGGAACTGTCT	ATCTGTTTTCAGACAAGAAAGCCAAACCTGGAGAGGAGCTTAAAGGTA	370
Qy	178	CTTAAGGAATATCT	GTTCGAGAACTAAATCCAAATGAACTGTACAGGCCCAATTGGAA	237
Db	371	CTGAAGGAATCTCT	GTTCGGGAATTTAAATCCAAATGAGACCGTGCAGGCCCAATGTGGAA	430
Qy	238	GCCCAACTCCTCT	CCCAAGCTGGACCAACCAGGCATTGTCAAGTTCATCAAGTTTGTG	297

Db	431	GCTCAGCTCTCTCCAGGCTCCATCATCTCTGCGCATGTGTCAGATTCCATGCAAGCTTCATG	490
Qy	298	GAGCAAGATAAATTTCTGCATTTATCAGCGAGTACTGTGAGGGCCGAGATCTGACGACATAAA	357
Db	491	GACGAGGAGACTTTTGTGCTATATCAGGAATACTGTGAGGGCCGAGATCTGACTATAGA	550
Qy	358	ATTGAGGATATTAACAGCTGGAATAATCTTCCAGAAAATCAATATATAGAAATGGTTT	417
Db	551	ATCCAGGAATATAAAGAGCTGGGAAGTCTTCGCTGAAAATCAGATAGATAAGTGGTTT	610
Qy	418	ATCCAGCTGCTGCTGGAGTTGACTACATGCATGAGAGGAGATCTTTCATCCAGACTTA	477
Db	611	ATCCAGCTGTTGCTGGAGTTGACTACATGCATGAGAGGAGATCTTTCATCCAGACTTG	670
Qy	478	AATCAAGAAATATATTTCTGAAAATTAATCTCCCTTAAATTTGGAGATTTTGGATTCT	537
Db	671	AAATCAAGAAATATATTTCTGAAAATTAATCTACTCAAAATCGGGATTTTGGAGTTCT	730
Qy	538	CGACTTCTAATGGATCTCTGACCTGGCCACAACTTTAACTGGAATCTCCCATATATG	597
Db	731	CGGCTGCTAATGGTTTCATGTGAGCTGGCTACAACTCTAACCGGACCCCCCATATATG	790
Qy	598	AGTCTGAGGCTCTGAAACACCAAGGCTATGACACAAAATGCGACATCTGGTCACTGGCA	657
Db	791	AGTCCGAGGCCCTGAAGACCAAGGCTATGATGCCAAGTCTGACATCTGGTCACTGGCA	850
Qy	658	TGCATTTTGTATGATGTGCTGCATGATCATGCAATTCGCTGGCTCCCAATTTCTTTATCC	717
Db	851	TGCATTTTATGATGTGCTGCATGATCATGCAATTCGCTGGCTCCCAATTTCTTTATCC	910
Qy	718	ATTGTTTAAAAATTTGTTGAAGTGTGACACACCTTCTCTCCCTGAGAGATATCCAAAGAA	777
Db	911	GTGTTTGTATTTGTTGAAGTGTGACACACCTTCTCTCCCTGAGAGATATCCAAAGAA	970
Qy	778	CTAATGCCATCATGGAAGCATGTTGAAACAGAAATCTTCTTCTTAAAGACATCTGCTATC	837
Db	971	CTAATGCCATCATGGAAGCATGTTGAAACAGAAATCTTCTTCTTAAAGACATCTGCTATC	1030
Qy	838	GAATTTTAAAAATCCCTTCTTGTATGAGCAGCTACAGACCTTAATGTGTAGATATCCA	897
Db	1031	GACATTTTAAAGCCCTTACATGGAAGCAGCTTCTGCTCTGATGTGTAATACCCA	1090
Qy	898	GAATGACTCTGGAAGCAAAAATTTGGATGTGTCAGAGAGGCTCTCATATTAATAT	957
Db	1091	GAGATGACTCTGGAAGCAAAAATTTGGATGTGTCAGAGAGGCTCTCATATTAATAT	1150
Qy	958	GCCATGCAAAAAGGATCCACCTGCGACACTCTGAGGCACTGTGCAAGTACAGAAAATG	1017
Db	1151	GCGGTGCAAAAAGGATCCACCTGCGACACTCTGAGGCACTGTGCAAGTACAGAAAATG	1210
Qy	1018	ACGCCAAGAGAAGGATCGCGCTGAGGAAGCTCCAGGCGCTCATGAGAAAGCCAGGAAG	1077
Db	1211	ACTCCAGAGAAACGATGTGGCTGAGGAAGCTGCGAGCAGCCGATCAGAGAGGAGG	1270
Qy	1078	CTGAAAAGATTTGCGAAGAAATATGAGAAATATGAAAGATGCAAGATTCGAGAAATGAGA	1137
Db	1271	CTGAAAAGATTTGCGAAGAAATATGAGAAATATGAAAGATGCAAGATTCGAGAAATGAGA	1330
Qy	1138	TCTCGGAATTTGCGAAGATGAGTGTGATGTGATGTACTCCATGAA	1179
Db	1331	TCCCGGAATGTTGGGTCGCTGATGCCATGCTGCTCCATGAA	1372
RESULT 3			
AK043543			
LOCUS	1784 bp mRNA linear HTC 19-SEP-2003		
DEFINITION	Mus musculus 10 days neonate cortex cDNA, RIKEN full-length enriched library, clone: A830007B13 product: similar to HUMAN HYPOPHYSAL PROTEIN FLJ23495, COMPLETE CDS (Macaca fascicularis), full insert sequence.		
ACCESSION	AK043543		
VERSION	AK043543.1 GI:26335750		
KEYWORDS	HTC; CAP trapper.		

SOURCE	Mus musculus (house mouse)
ORGANISM	Mus musculus
REFERENCE	Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.
AUTHORS	Carninci, P. and Hayashizaki, Y.
TITLE	High-efficiency full-length cDNA cloning
JOURNAL	Meth. Enzymol. 303, 19-44 (1999)
MEDLINE	99279253
PUBMED	10349636
REFERENCE	Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K., Itoh, M., Konno, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y. Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes Genome Res. 10 (10), 1617-1630 (2000)
AUTHORS	20499374
TITLE	11042159
JOURNAL	Shibata, K., Itoh, M., Aizawa, K., Nagaoka, S., Sasaki, N., Carninci, P., Konno, H., Akiyama, J., Nishi, K., Kishimoto, T., Tashiro, H., Itoh, M., Sumi, N., Ishii, Y., Nakamura, S., Hazama, M., Nishine, T., Hatake, A., Yamamoto, R., Matsumoto, H., Sakaguchi, S., Ikegami, T., Kashiwagi, K., Fujiwara, S., Inoue, K., Togawa, Y., Izawa, M., Chata, E., Watanabe, M., Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsuura, S., Kawai, O., Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A. and Hayashizaki, Y. RIKEN integrated sequence analysis (RISA) system--384-format sequencing pipeline with 384 multicapillary sequencer Genome Res. 10 (11), 1757-1771 (2000)
AUTHORS	2050913
TITLE	11076861
JOURNAL	The RIKEN Genome Exploration Research Group Phase II Team and the FANTOM Consortium.
MEDLINE	Functional annotation of a full-length mouse cDNA collection Nature 409, 685-690 (2001)
PUBMED	5
REFERENCE	The FANTOM Consortium and the RIKEN Genome Exploration Research Group Phase I & II Team. Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs Nature 420, 563-573 (2002)
AUTHORS	6 (bases 1 to 1784)
TITLE	Adachi, J., Aizawa, K., Akimura, T., Arakawa, T., Bono, H., Carninci, P., Fukuda, S., Furuno, M., Hanagaki, T., Hara, A., Hashizume, W., Hayashida, K., Hayatsu, N., Hiramoto, K., Hiraoka, I., Hirozane, T., Hori, F., Imotani, K., Ishii, Y., Itoh, M., Kagawa, I., Kasukawa, T., Katoh, H., Kawai, J., Kojima, Y., Kondo, S., Konno, H., Kouda, M., Koya, S., Kurihara, C., Matsuyama, T., Miyazaki, A., Murata, M., Nakamura, M., Nishi, K., Nomura, K., Numazaki, R., Ohno, M., Ohsato, N., Okazaki, Y., Saitoh, R., Saitoh, H., Sakai, C., Sakai, K., Sakazume, N., Sano, H., Sasaki, D., Shibata, K., Shinagawa, A., Shiraki, T., Sobue, Y., Tagami, M., Tagawa, A., Takahashi, F., Takaku-Akahira, S., Takeda, Y., Tanaka, T., Tomaru, A., Toya, T., Yasunishi, A., Muramatsu, M. and Hayashizaki, Y. Direct Submission
AUTHORS	Submitted (15-JUL-2001) Yoshihide Hayashizaki, The Institute of Physical and Chemical Research (RIKEN) Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center (GSC), RIKEN Yokohama Institute; 1-7-22 Suehiro-cho, Tsurumi-Ku, Yokohama, Kanagawa 230-0045, Japan (E-mail: genome-res@gsc.riken.go.jp, URL: http://genome.gsc.riken.go.jp/, Tel: 81-45-503-9222, Fax: 81-45-503-9216)
JOURNAL	cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in RIKEN Division of Experimental Animal Research in Riken contributed to prepare mouse tissues.
REFERENCE	Please visit our web site for further details. URL: http://genome.gsc.riken.go.jp/ URL: http://fantom.gsc.riken.go.jp/ Location/Qualifiers
FEATURES	1. ..1784
source	





<div><div>/db_xref="taxon:9606" /clone="CS0DJ004YB13" /cell_type="T CELLS (JURKAT CELL LINE) COT 10-NORMALIZED" /cell_line="JURKAT" /clone_lib="Homo sapiens T CELLS (JURKAT CELL LINE) COT 10-NORMALIZED" /notes="1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime end enriched, double-strand cDNA was digested with Not I and EcoR V and cloned into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library was normalized."</div></div>									
ORIGIN									
Query Match 58.1%; Score 841.4; DB 13; Length 1201;									
Best Local Similarity 96.0%; Pred. No. 4.5e-190;									
Matches 882; Conservative 12; Mismatches 18; Indels 7; Gaps 3;									
QY	1	ATGCTGAAATTC	CAAGAGG	CAGCTAAGTGTGTGAGTGGATCAACAGCCATTTCCACTTAT	60				
DB	290	ATGCTGAAATTC	CAAGAGG	CAGCTAAGTGTGTGAGTGGATCAACAGCCATTTCCACTTAT	349				
QY	61	CCAAAGACCTTG	ATTCGAAGA	GATACGTGCTTCAACAAAACTTCGCGAGTGAAGTTTT	120				
DB	350	CCAAAGACCTTG	ATTCGAAGA	GATACGTGCTTCAACAAAACTTCGCGAGTGAAGTTTT	409				
QY	121	GGAACTGTCTAT	CTGCTGTTT	CAGACAAGAAAGCCAAACGAGGACGAGAAATTAAGGTACTT	180				
DB	410	GGAACTGTCTAT	CTGCTGTTT	CAGACAAGAAAGCCAAACGAGGACGAGAAATTAAGGTACTT	469				
QY	181	AAGGAAATATCT	TTTGGAGAACT	TAATTCCAAATGAAACTGTACAGGCCAATTTGGAAGCC	240				
DB	470	AAGGAAATATCT	TTTGGAGAACT	TAATTCCAAATGAAACTGTACAGGCCAATTTGGAAGCC	529				
QY	241	CRACCTCTCTC	CAAGCTGGAC	CAACCCAGCCATTTGCTCAAGTTCATCGATCGAAGTTTGTGGAG	300				
DB	530	CRACCTCTCTC	CAAGCTGGAC	CAACCCAGCCATTTGCTCAAGTTCATCGATCGAAGTTTGTGGAG	589				
QY	301	CAAGATAAATTT	CTGCAATAT	CACGGAGTACTGTGAGGGCGGAGATCTGACGCAATAAAAT	360				
DB	590	CAAGATAAATTT	CTGCAATAT	CACGGAGTACTGTGAGGGCGGAGATCTGACGCAATAAAAT	649				
QY	361	CAGGAATATAA	CAAGCTG	GAATAATCTTCCAGAAATCAAAATATAGATGGTTATC	420				
DB	650	CAGGAATATAA	CAAGCTG	GAATAATCTTCCAGAAATCAAAATATAGATGGTTATC	709				
QY	421	CAGCTGCTG	CTGGGAGTTG	ACTACATGCATGAGAGGAGATCTTCATCGAGACTTTAAAG	480				
DB	710	CAGCTGCTG	CTGGGAGTTG	ACTACATGCATGAGAGGAGATCTTCATCGAGACTTTAAAG	769				
QY	481	TCAAAGAAATG	TATTTCTG	AAAAATAATCTCTTTAAATTTGGAGATTTTGGAGTTTCTCGA	540				
DB	770	TCAAAGAAATG	TATTTCTG	AAAAATAATCTCTTTAAATTTGGAGATTTTGGAGTTTCTCGA	829				
QY	541	CTTCTAATGG	ATCCTGTG	ACCTGGCCACAACCTTTAACTGGAACTCCCAATATATAGAT	600				
DB	830	CTTCTAATGG	ATCCTGTG	ACCTGGCCACAACCTTTAACTGGAACTCCCAATATATAGAT	889				
QY	601	CTTGAGGCTCT	GAAACAC	CAAGGCTATGACAAAGTGGGACATCTGGTCACTGGCATGC	660				
DB	890	CTTGAGGCTCT	GAAACAC	CAAGGCTATGACAAAGTGGGACATCTGGTCACTGGCATGC	949				
QY	661	ATTTTCTATG	ATGATGCT	GCATGAATTCATGCATTCGCTGGCTCCAAATTTCTTTATCCATT	720				
DB	950	ATTTTCTATG	ATGATGCT	GCATGAATTCATGCATTCGCTGGCTCCAAATTTCTTTATCCATT	1009				
QY	721	GTTTTAAAAAT	TGTTGAAG	TGACACACCTTCTCTCCCTGAGAGATATCCAAAGAACTA	780				
DB	1010	GTTTTAAAAAT	TGTTGAAG	TGACACACCTTCTCTCCCTGAGAGAAATCCAAAGAACTA	1069				
QY	781	AATGCCATCAT	GGAAGCATG	TGTAACAGAAATCCCTTCAATTAAGACCATCTGCTATCGAA	840				
DB	1070	AATGCCATCAT	GGAAGCATG	TGTAACAGAAATCCCTTCAATTAAGACCATCTGCTATCGAA	1128				
QY	841	ATTTTAAAAAT	TCCTTTAC	CTTTGATGAGCAGCTACAGAACCTTAATGTGTAGATATTCAGAA	900				

```

Db      1129 WTTTAAAYCYCTTCCTGTGAGCGC----ACAGAACCCTAATGTGTGK--WTTTCWGA 1182
QY      901 ATGACTCTGGAGACAAAA 919
Db      1183 AAGACTCTGRRGACAAA 1201

RESULT 5
AK054237
LOCUS   1221 bp mRNA linear HTC 20-SEP-2003
DEFINITION Mus musculus 2 days pregnant adult female oviduct cDNA, RIKEN full-length enriched library, clone:E230038G21 product:similar to mRNA, SIMILAR TO HUMAN HYPOTHETICAL PROTEIN FLJ23495, COMPLETE CDS [Macaca fascicularis], full insert sequence.
ACCESSION AK054237
VERSION AK054237.1 GI:26344084
KEYWORDS HTC; CAP trapper.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sclurognathi; Muridae; Murinae; Mus.
1
REFERENCE
AUTHORS Carninci,P. and Hayashizaki,Y.
TITLE High-efficiency full-length cDNA cloning
JOURNAL Meth. Enzymol. 303, 19-44 (1999)
MEDLINE 99279253
PubMed 10349636
2
REFERENCE
AUTHORS Carninci,P., Shibata,Y., Hayatsu,N., Sugahara,Y., Shibata,K., Itoh,M., Konno,H., Okazaki,Y., Muramatsu,M. and Hayashizaki,Y.
TITLE Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes
JOURNAL Genome Res. 10 (10), 1617-1630 (2000)
MEDLINE 20499374
PubMed 11042159
3
REFERENCE
AUTHORS Shibata,K., Itoh,M., Aizawa,K., Nagaoka,S., Sasaki,N., Carninci,P., Konno,H., Akiyama,J., Nishi,K., Kitsuai,T., Teshiro,H., Itoh,M., Sumi,N., Ishii,Y., Nakamura,S., Hazama,M., Nishine,T., Harada,A., Yamamoto,R., Matsumoto,H., Sakaguchi,S., Ikegami,T., Kashiwagi,K., Yoneda,Y., Ishikawa,T., Ogawa,K., Tanaka,T., Matsura,S., Kawai,J., Okazaki,Y., Muramatsu,M., Inoue,Y., Kira,A. and Hayashizaki,Y.
TITLE RIKEN integrated sequence analysis (RISA) system--384-format sequencing pipeline with 384 multicapillary sequencer
JOURNAL Genome Res. 10 (11), 1757-1771 (2000)
MEDLINE 20530913
PubMed 11076861
4
REFERENCE
AUTHORS The RIKEN Genome Exploration Research Group Phase II Team and the FANTOM Consortium.
TITLE Functional annotation of a full-length mouse cDNA collection
JOURNAL Nature 409, 695-690 (2001)
5
REFERENCE
AUTHORS The FANTOM Consortium and the RIKEN Genome Exploration Research Group Phase I & II Team.
TITLE Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs
JOURNAL Nature 420, 563-573 (2002)
6 (bases 1 to 1221)
REFERENCE
AUTHORS Adachi,U., Aizawa,K., Akimura,T., Arakawa,T., Bono,H., Carninci,P., Fukuda,S., Furuno,M., Hanagaki,T., Hara,A., Hashizume,W., Hayashida,K., Hayatsu,N., Hiramoto,K., Hiroaka,T., Hirozane,T., Hori,F., Imotani,K., Ishii,Y., Itoh,M., Kagawa,I., Kasukawa,T., Katoh,H., Kawai,J., Kojima,Y., Kondo,S., Konno,H., Kouda,M., Koya,S., Kurihara,C., Matsuyama,T., Miyazaki,A., Murata,M., Nakamura,M., Nishi,K., Nomura,K., Numazaki,R., Ohno,M., Ohsato,N., Okazaki,Y., Saito,R., Saich,H., Sakai,C., Sakai,K., Sakazume,N., Sano,H., Sasaki,D., Shibata,K., Shingawa,A., Shiraki,T., Scoble,Y., Tagami,M., Tagawa,K., Takahashi,F., Takaku-Akahira,S., Takeda,Y., Tanaka,T., Tomaru,A., Toya,T., Yasunishi,A., Muramatsu,M. and Hayashizaki,Y.

```

TITLE  
JOURNAL

Direct Submission  
Submitted (16-JUL-2001) Yoshihide Hayashizaki, The Institute of Physical and Chemical Research (RIKEN), Laboratory for Genome Exploration and Research Group, RIKEN Genomic Sciences Center (GSC), RIKEN Yokohama Institute, 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan (E-mail: genome-res@sc.riken.go.jp, URL: http://genome.gsc.riken.go.jp/, Tel: 81-45-503-9222, Fax: 81-45-503-9216)

COMMENT  
cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in RIKEN. Division of Experimental Animal Research in Riken contributed to prepare mouse tissues.  
Please visit our web site for further details.  
URL: http://genome.gsc.riken.go.jp/  
URL: http://fantom.gsc.riken.go.jp/  
Location/Qualifiers  
1. .1221  
/organism="Mus musculus"  
/mol\_type="mRNA"  
/strain="C57BL/6J"  
/db\_xref="PANTOM DB:E230038G21"  
/db\_xref="MGI:2426244"  
/db\_xref="taxon:10090"  
/clone="E230038G21"  
/sex="female"  
/tissue type="oviduct"  
/clone\_lib="RIKEN full-length enriched mouse cDNA library"  
/dev\_stage="2 days pregnant adult"  
126. .>1220  
/note="unnamed protein product; putative similar to MRNA, SIMILAR TO HUMAN HYPOTHETICAL PROTEIN FLJ23495, COMPLETE CDS [Macaca fascicularis] (SPTRI|BAB83539, evidence: FASTY, 72.4%ID, 100%length, match=1806)"  
/codon\_start=1  
/protein\_id="BAC35699.1"  
/db\_xref="GI:26344085"  
/translation="MLKQSTAKVGRRTVPMYPTALIRYVYLQKLGSGFTVYVSKKAKGDELIVKISVGEINPNETVOANVEAQLSLHHPAIVRPHASFMQEFPCITTECEGRDIDRTOEYKEAGKFAENOIVFQFTOLLGVYVHERRILRDLKSNIFLKNLAKIGDPGVSRLMGSCELATLTGTPTPHYSPEALKHOGYDAKSDIWSLACILYEMCLOHAFAGSSFLSVNLIVGKTPSLPDYPRELNTIMERNLKSRLRPSAAILKAPIMEEQQLWCKPTEMLDKNSVCQKEAAHTINAVQKHLHLQLOALSDTKTTPPRMWRMLKQADDERARLKKIA"

CDS

Query Match 55.9%; Score 809.8; DB 11; Length 1221;  
Best Local Similarity 84.4%; Pred. No. 1 6e-182;  
Matches 923; Conservative 0; Mismatches 167; Indels 3; Gaps 1;

QY 1 ATGCTGAATTCAGAGGAGCTAAAGTGTGTG---AGTGGATCAACAGCCATTTCCACT 57  
126 ATGCTGAATTCAGAGGAGCTGGAAGTGTGTGGCAGAGAGCCACAGTCATCCCCATG 185

QY 58 TATCCAAAGACCTTGATTCAGAGAGTACGTCCTTCAACAAAATCTTGCAGTGGAGT 117  
Db 186 TACCCGACCGCTTGATTCAGAGAGATATGTCCTTCAGAGAAATCTTGCAGTGGAGT 245

QY 118 TTGGAACTGTCTATCTGTTTCAGACAGAAAGCCAAACGAGGAGGAGAAATTAAGGTA 177  
Db 246 TTGGAACTGTCTATCTGTTTCAGACAGAGAAAGCCAAACCTGGAGAGGCTAAAGTA 305

QY 178 CTTAAGGAAATATCTGTTGAGAACTTAATCCAAATGAAACTGTACAGGCCAATTGGAA 237  
Db 306 CTGAAGGAAATCTCTGTTGGCGAATTAATCCAAATGAGACCGTGCAGGCCAATGTGAA 365

QY 238 GCCCACTCTCTCCAAAGCTGGACCCAGCCATGTCAGTTCCTCAAGTTCCTCAAGTTTGTG 297  
Db 366 GCTCAGCTCTCTCCAGGCTCCATCATCTGCGCATTTGTGATTCATCAAGCTTCATG 425

QY 298 GACCAAGATAATTTCTGCATTATCAACGAGTACTGTGAGGGCCGAGATCTGACGAGTAA 357

Db 426 GAGCAGGAGACCTTTTGCATTATCAGCGAATACTGTGAGGSCCGAGATCTCGACTATAGA 485  
QY 358 ATTACGGAATTAACAAGCTGGAAAATCTTTCAGAAAATCAAAATTAATAGATGTTT 417  
Db 486 ATCCAGGAATTAAGAAGAGCTGGGAAAATCTTTCGCTGAAAATCAGATAGTAAATGTTT 545  
QY 418 ATCCAGCTGCTGCTGGGAGTTGACTACATGATCAGAGGAGGAGTACTTTCATCGAGACTTA 477  
Db 546 ATCCAGCTGCTGCTGGGAGTTGACTACATGATCAGAGGAGGAGTACTTTCATCGAGACTTA 605  
QY 478 AAGTCAAGAAATGATTTTCTCAAAAATAATCTCTTAAATTTGAGAGATTTTGGATTTCT 537  
Db 606 AAATCAAGAAATATATTTCTGAAAATAATCTTACTCAAAATCGGGGATTTTGGAGTTTCT 665  
QY 538 CGACTTCTAATGGGATCTGCTGACCTGGCCACAACTTTAACTGGAACCTCCCATTTATG 597  
Db 666 CGGCTGCTAATGGGTTCAITGAGCTGGCTAACTCTAACCGGAGCCCCCATTTATG 725  
QY 598 AGTCTGAGGCTCTGAAACACCAAGGCTATGACACAAAGTCGACATCTGGTCACTGGCA 657  
Db 756 AGTCCGAGGCGCTGAAGCACCAGGCTATGATGCCAAGTCTGACATCTGGTCACTGGCA 785  
QY 658 TGCATTTTGTATGATGATGCTGCTGCATGATGATGATGATGATGATGATGATGATGATG 717  
Db 786 TGCATTTTATATGATGATGCTGCTGCATGATGATGATGATGATGATGATGATGATGATG 845  
QY 718 ATTGTTTTAAAAATTTGTAAGGTCACACACCTTCTCTCCCTGAGAGATATCCAAAAGAA 777  
Db 846 GTGGTTTGTATATTTGTAAGGTTAAACACCTTCTCTCCCGGACAGATATCCACGAGAA 905  
QY 778 CTAATGTCATCATGGAAGCATGTTGAAAGAAATCTTTCATTAAGACCATCTGCTATC 837  
Db 906 CTAACACCATCATGGAAGCATGTTGAAAGAAATCTTTCATTAAGACCATCTGCTATC 965  
QY 838 GAAATTTTAAAAATCCCTTACCTTGTGAGCAGCTACAGAACCTTAATGTGTAGATATTCA 897  
Db 966 GACATTTTAAAAAGCCCTTATCATGGAAGAGCAGCTTCACTCTCTGATGTGTAAATACCCA 1025  
QY 898 GAAATGCTCTGGAAGCAAAAATTTGGATTTGTCAGAGGAGGCTGCTCATATTAATTAAT 957  
Db 1026 GAGATGACACTGGAAGCAAGAACTCAGTTTGTGTCAGAGGAGGCTGCTCATACCATTAAT 1085  
QY 958 GCCATGCAAAAAGAGTCCACTCGACAGCTCGAGGACCTGTGAGAGTACAGAAATG 1017  
Db 1086 GCCGTGCAAAAAGAGTTCACCTGACACTCTGCAAGCTTGTCTGACCCAGAGAAACG 1145  
QY 1018 ACGCCAAGAGAAAGATGCGGCTGAGGAAGCTCCAGGCGGCTGATGAGAAACCCAGGAAG 1077  
Db 1146 ACTCCAGAGAAACGAGTGTGCTGAGGAAGCTGCGAGGAGCCGATGAGAGAGCCAGGAGG 1205  
QY 1078 CTGAAAAGATTG 1090  
Db 1206 CTGAAAAGATTG 1218

RESULT 6  
BQ709169  
LOCUS  
DEFINITION  
BQ709169  
ACCESSION  
VERSION  
KEYWORDS  
SOURCE  
ORGANISM  
REFERENCE  
AUTHORS  
TITLE  
JOURNAL  
COMMENT

BQ709169 937 bp mRNA linear EST 16-JUL-2002  
AGENCOURT\_8495234 NIH\_MGC\_113 Homo sapiens cDNA clone IMAGE:6301955  
5', mRNA sequence.  
BQ709169  
BQ709169.1 GI:21848068  
EST  
Homo sapiens (human)  
Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
1 (bases 1 to 937)  
NIH-MGC http://mgi.nci.nih.gov/  
National Institutes of Health, Mammalian Gene Collection (MGC)  
Unpublished (1999)  
Contact: Robert Strausberg, Ph.D.  
Email: cgapbs-remail.nih.gov

Tissue Procurement: Dr. Mark Watson  
cDNA Library Preparation: Rubin Laboratory  
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)  
DNA Sequencing by: Agencourt Bioscience Corporation  
Clone distribution: MGC clone distribution information can be  
found through the I.M.A.G.E. Consortium/LLNL at:  
http://image.llnl.gov

Plate: L1CM2518 row: k column: 12  
High quality sequence start: 4  
High quality sequence stop: 638.

## FEATURES

source

1..937  
Location/Qualifiers  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="IMAGE:6301955"  
/lab\_host="DH10B (phage-resistant)"  
/clone\_lib="NIH\_MGC\_113"  
/note="Organ: spleen; Vector: pOT37; Site\_1: XhoI; Site\_2:  
EcoRI; cDNA made by oligo-dT priming. Directionally cloned  
into EcoRI/XhoI sites using the following 5' adaptor:  
GGCAGCAG(G). Library constructed by Ling Hong in the  
laboratory of Gerald M. Rubin (University of California,  
Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and  
Superscript II RT (Life Technologies). Note: this is a  
NIH\_MGC Library."

## ORIGIN

Query Match 49.3%; Score 714.4; DB 13; Length 937;  
Best Local Similarity 96.9%; Pred. No. 9.2e-160;  
Matches 760; Conservative 0; Mismatches 21; Indels 3; Gaps 3;  
QY 1 ATGCTGAATTCACAGGAGGAGTAAAGTGTGAGTGGATCAACAGCCATTTCCATTAT 60  
DB 153 ATGCTGAATTCACAGGAGGAGTAAAGTGTGAGTGGATCAACAGCCATTTCCATTAT 212  
QY 61 CCAAGACCTTGATTCGACAGAGATAGTCTCTTCAACAAAACTTGGCAGTGAAGTTT 120  
DB 213 CCAAGACCTTGATTCGACAGAGATAGTCTCTTCAACAAAACTTGGCAGTGAAGTTT 272  
QY 121 GGAACCTGTATCTGTTTCAGACAGAAAGCCAAACGAGGAGAGGAATTAAGTACTT 180  
DB 273 GGAACCTGTATCTGTTTCAGACAGAAAGCCAAACGAGGAGGAATTAAGTACTT 332  
QY 181 AAGGAATATCTGTGTGAGAACTTAATCCAAATGAAATGTACAGGCCAATTTGGAAGCC 240  
DB 333 AAGGAATATCTGTGTGAGAACTTAATCCAAATGAAATGTACAGGCCAATTTGGAAGCC 392  
QY 241 CAACCTCTCCAGAGTGGACCCAGCCAGCCATTTCAAGTTCCATCAAGTTTGTGGAG 300  
DB 393 CAACCTCTCCAGAGTGGACCCAGCCAGCCATTTCAAGTTCCATCAAGTTTGTGGAG 452  
QY 301 CAAGATAATTTCTGCATTATCAGGAGTACTGTGAGGCCGAGATCTGCAGCATAAAT 360  
DB 453 CAAGATAATTTCTGCATTATCAGGAGTACTGTGAGGCCGAGATCTGCAGCATAAAT 512  
QY 361 CAGGAATATAAACAAGCTGGAATAATCTTTCCAGAAAAATCAATAATAGAAATGTTTATC 420  
DB 513 CAGGAATATAAACAAGCTGGAATAATCTTTCCAGAAAAATCAATAATAGAAATGTTTATC 572  
QY 421 CAGCTGCTGGAGTTGATCATGATGAGAGGAGATCTTCAATCGAGACTTAAG 480  
DB 573 CAGCTGCTGGAGTTGATCATGATGAGAGGAGATCTTCAATCGAGACTTAAG 632  
QY 481 TCAAGAAATCTATTCTGAAAAATATCTCTTAAAAATGGAGATTTGGAGTTTCTCGA 540  
DB 633 TCAAGAAATCTATTCTGAAAAATATCTCTTAAAAATGGAGATTTGGAGTTTCTCGA 692  
QY 541 CTTCTAATGGAT-CTGTGACCTGGGACCAACATTTTAACCTGGAACTCCCAATATATGAG 599  
DB 693 CTTCTAATGGATCCCTGTGACCTGGGACCAACATTTTAACCTGGAACTCCCAATATATGAG 752  
QY 600 TCCTGAGGCTCTGAACACCAAGGCTATGACACAAAGTCGACATCT-GGTCACTGGCAT 658

DB 753 TCCTGAGGCTCTGAACCCCAAGGCTATGACACAAAGTCGACATCTGGGTCACATG 812  
QY 659 GCATTTTGTATGAGATGTGTCATGAATCATGATTCGCTGGCTCCAAATTTCTATCCA 718  
DB 813 GCATTTTGTATGAGATGTGTCATGAATCATGATTCGCTGGCTCCAAATTTCTATCCA 872  
QY 719 TGTGTTAAAAAATGTTGAAGGTGACA-CACCTTCTCTCCTCTGAGATATCCAAAGAA 777  
DB 873 ATGTTTAAAAAATGTTGAAGGGGACACCACTTCCCTCCCTGGGAAATATTCAAAGAA 932  
QY 778 CTAA 781  
DB 933 CTAA 936

## RESULT 7

BX328072

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

BX328072 932 bp mRNA linear EST 01-MAY-2003  
BX328072 Homo sapiens T CELLS (JURKAT CELL LINE) COT 10-NORMALIZED  
Homo sapiens cDNA clone CS00J004YB13 5-PRIME, mRNA sequence.

BX328072

EST.

Homo sapiens (human)

Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

1. (bases 1 to 932)

Li.W.B., Gruber,C., Jesse,J. and Polayes,D.

Full-length cDNA libraries and normalization

Unpublished (2001)

Contact: Genoscope

Genoscope - Centre National de Sequencage

BP 191 91006 Evry cedex - France

Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr

Library was constructed by Life Technologies, a division of

Invitrogen. This sequence belongs to sequence cluster 2479.f for

more information about this cluster, see

http://www.genoscope.cns.fr/

cgi-bin/cluster.cgi?seq=CS0BAG041ZF11\_CS03926\_1acluster=2479.f.

Contact : Feng Liang Email : fliang@lifetech.com URL :

http://fulllength.invitrogen.com/Invitrogen Corporation 1600

Faraday Avenue Genoscope sequence ID : CS0BAG041ZF11\_CS03926\_1.

Location/Qualifiers

1..932

/organism="Homo sapiens"

/mol\_type="mRNA"

/db\_xref="taxon:9606"

/clone="CS00J004YB13"

/cell\_type="T CELLS (JURKAT CELL LINE) COT 10-NORMALIZED"

/cell\_line="JURKAT"

/clone\_lib="Homo sapiens T CELLS (JURKAT CELL LINE) COT

10-NORMALIZED"

/note="1st strand cDNA was primed with a NotI-oligo(dT)

primer. Five prime end enriched, double-strand cDNA was

digested with Not I and cloned into the Not I and EcoR V

sites of the pCMVSPORT 6 vector. Library was normalized."

ORIGIN

Query Match 47.8%; Score 692.4; DB 13; Length 932;

Best Local Similarity 98.7%; Pred. No. 1.7e-154;

Matches 708; Conservative 0; Mismatches 7; Indels 2; Gaps 1;

QY 683 TGAATCATGCAATCGCTGGCTCCAAATTTCTTATCCATTTGTTTAAATAATTTGTTGAAGTG 742

DB 11 TGNATCATGCAATCGCTGGCTCCA--TTTCTTATCTCTTTGTTTAAATAATTTGTTGAAGTG 68

QY 743 ACACACCTTCTCTCCCTGAGAGATATCCAAAGAACTTAATGATCCATCATCGAAGCATGT 802

DB 69 ACACACCTTCTCTCCCTGAGAGATATCCAAAGAACTTAATGATCCATCATCGAAGCATGT 128

QY 803 TGAACAAGATCTCTTATTAAGACCATCTGTCGAAATTTTAAATAATTTTACCTTG 862

```

Db      129  TGAACAGAACTCTTCATTAAAGACCATCTGCTATCGAAATTTTAAATAACCTTACCTTG 188
QY      863  ATGAGCAGCTACAGACCTTAATGTGTAGATATTAGAAATGACTCTGGAGACAAAAATT 922
Db      189  ATGAGCAGCTACAGACCTTAATGTGTAGATATTAGAAATGACTCTGGAGACAAAAATT 248
QY      923  TGGATTGTTCAGAGGAGGCTGCTCATATTAATTAATCCATGCAAAAAGATCCACCTGC 982
Db      249  TGGATTGTTCAGAGGAGGCTGCTCATATTAATTAATCCATGCAAAAAGATCCACCTGC 308
QY      983  AGACTCTGAGGGCACTGTTCAGAGTACAGAAATGACGCAAGAGAGAAAGATGCGGCTGA 1042
Db      309  AGACTCTGAGGGCACTGTTCAGAGTACAGAAATGACGCAAGAGAGAAAGATGCGGCTGA 368
QY      1043  GGAAGCTCCAGGGCGGTGTGAGAAAGCCAGGAGCTGAAAAGATTTGGAGAGAAAT 1102
Db      369  GGAAGCTCCAGGGCGGTGTGAGAAAGCCAGGAGCTGAAAAGATTTGGAGAGAAAT 428
QY      1103  ATGAAGAAATAGCAAAACGAATCCAGAAATTTGAGATCTCGGAACCTTTCAGACGCTGAGTG 1162
Db      429  ATGAAGAAATAGCAAAACGAATCCAGAAATTTGAGATCTCGGAACCTTTCAGACGCTGAGTG 488
QY      1163  TTGATCTACTCCATGAAAACACATTTAAAGGAATGGAAGAAAGAGGAGGACCACTG 1222
Db      489  TTGATCTACTCCATGAAAACACATTTAAAGGAATGGAAGAAAGAGGAGGACCACTG 548
QY      1223  AGGGAAGACTTCTTGTTCACCCAGGAGGAGATGAGAGAGGAGGAGGAGGAGGAGGAG 1282
Db      549  AGGGAAGACTTCTTGTTCACCCAGGAGGAGATGAGAGAGGAGGAGGAGGAGGAGGAG 608
QY      1283  AGGAATCTGATGAAACCACTTTAGAGAACTGCTGAGTCTGAGCTATTCCTTCATGCG 1342
Db      609  AGGAATCTGATGAAACCACTTTAGAGAACTGCTGAGTCTGAGCTATTCCTTCATGCG 668
QY      1343  ACCTCCAGCAACTTGAATCAATTTGAGAGATGCCACATCTGACCTTGGATACCATG 1399
Db      669  ACCTCCAGCAACTTGAATCAATTTGAGAGATGCCACATCTGACCTTGGATACCATG 725

RESULT 8
AK030186      2239 bp  mRNA  linear  HTC 18-SPB-2003
LOCUS      Mus musculus adult male testis cDNA, RIKEN full-length enriched
DEFINITION library, clone:4933413H15 product:similar to MRNA, SIMILAR TO HUMAN
HYPOTHETICAL PROTEIN FL023495, COMPLETE CDS [Macaca fascicularis],
full insert sequence.
ACCESSION AK030186
VERSION AK030186.1 GI:26081701
KEYWORDS HTC; CAP trapper.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
REFERENCE 1
AUTHORS Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K.,
TITLE Itoh, M., Konno, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.
JOURNAL Normalization and subtraction of cap-trapper-selected cDNAs to
MEDLINE Prepare full-length cDNA libraries for rapid discovery of new genes
PUBMED Genome Res. 10 (10), 1617-1630 (2000)
REFERENCE 2
AUTHORS Shibata, K., Itoh, M., Aizawa, K., Nagaoka, S., Sasaki, N., Carninci, P.,
TITLE Konno, H., Akiyama, J., Nishi, K., Katsunai, T., Tashiro, H., Itoh, M.,
JOURNAL Sumi, N., Ishii, Y., Nakamura, S., Hazama, M., Nishine, I., Harada, A.,
MEDLINE Yamamoto, R., Matsumoto, H., Sakaguchi, S., Ikegami, T., Kashiwagi, K.,
PUBMED 11042159
REFERENCE 3
AUTHORS
TITLE
JOURNAL
MEDLINE
PUBMED
REFERENCE
AUTHORS
TITLE
JOURNAL
MEDLINE
PUBMED
REFERENCE
AUTHORS

```

```

Fujiwaka, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E., Watahiki, M.,
Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsuura, S., Kawai, J.,
Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A. and Hayashizaki, Y.
RIKEN integrated sequence analysis (RISA) system--384-format
sequencing pipeline with 384 multicapillary sequencer
Genome Res. 10 (11), 1757-1771 (2000)
20530913
MEDLINE
PUBMED 11076861
REFERENCE 4
AUTHORS The RIKEN Genome Exploration Research Group Phase II Team and the
FANTOM Consortium.
TITLE Functional annotation of a full-length mouse cDNA collection
JOURNAL Nature 409, 685-690 (2001)
REFERENCE 5
AUTHORS The FANTOM Consortium and the RIKEN Genome Exploration Research
Group Phase I & II Team.
TITLE Analysis of the mouse transcriptome based on functional annotation
of 60,770 full-length cDNAs
JOURNAL Nature 420, 563-573 (2002)
REFERENCE 6
AUTHORS Adachi, J., Aizawa, K., Akimura, T., Arakawa, T., Bono, H., Carninci, P.,
Fukuda, S., Furuno, M., Hanagaki, T., Hara, A., Hashizume, W.,
Hayashida, K., Hayatsu, N., Hiramoto, K., Hiraoka, T., Hirozane, T.,
Hori, F., Imocani, K., Ishii, Y., Itoh, M., Kagawa, I., Kasukawa, T.,
Kato, H., Kawai, J., Kojima, Y., Kondo, S., Konno, H., Kouda, M.,
Koya, S., Kurihara, C., Matsuyama, T., Miyazaki, A., Murata, M.,
Nakamura, N., Nishi, K., Nomura, K., Numazaki, R., Ohno, M., Ohsato, N.,
Okazaki, Y., Saitoh, H., Saitoh, H., Sakai, C., Sakai, K., Sakazume, N.,
Sano, H., Sasaki, D., Shibata, K., Shinagawa, A., Shiraki, T.,
Sogabe, Y., Tagami, M., Tagawa, A., Takahashi, P., Takaku-Akaira, S.,
Takeda, Y., Tanaka, I., Tomaru, A., Toyota, T., Yasunishi, A.,
Muramatsu, M. and Hayashizaki, Y.
Direct Submission
SUBMITTED (16-JUL-2001) Yoshihide Hayashizaki, The Institute of
Physical and Chemical Research (RIKEN), Laboratory for Genome
Exploration Research Group, RIKEN Genomic Sciences Center (GSC),
RIKEN Yokohama Institute, 1-7-22 Sushiro-cho, Tsurumi-ku, Yokohama,
Kanagawa 230-0045, Japan (E-mail:genome-res@gsc.riken.go.jp,
URL:http://genome.gsc.riken.go.jp/, Tel:81-45-503-9222,
Fax:81-45-503-9216)
cDNA library was prepared and sequenced in Mouse Genome
Encyclopedia Project of Genome Exploration Research Group in Riken
Genomic Sciences Center and Genome Science Laboratory in RIKEN.
Division of Experimental Animal Research in Riken contributed to
prepare mouse tissues.
Please visit our web site for further details.
URL:http://genome.gsc.riken.go.jp/
URL:http://fantom.gsc.riken.go.jp/.
FEATURES
Location/Qualifiers
source 1..2239
organism="Mus musculus"
mol_type="mRNA"
strain="C57BL/6J"
db_xref="FANTOM DB:4933413H15"
db_xref="MGI:2392882"
db_xref="taxon:10090"
clone="4933413H15"
sex="male"
tissue_type="testis"
clone_lib="RIKEN full-length enriched mouse cDNA library"
dev_stage="adult"
misc_feature 235..1598
note="putative
similar to MSNA, SIMILAR TO HUMAN HYPOTHETICAL PROTEIN
FL023495, COMPLETE CDS [Macaca fascicularis]
(SPR|BAB83539, evidence: PASTY, 72.4%ID, 100%length,
match=1806)"
polyA_signal 2221..2226
polyA_site 2239
note="putative"
ORIGIN

```

Query Match		42.7%; Score 618.8; DB 11; Length 2239;
Best Local Similarity		74.0%; Pred. No. 7e-137;
Matches 875; Conservative		0; Mismatches 177; Indels 130; Gaps 2;
Qy	1	ATGCTGAAATTCGAAGGAGCTAAGTGTGTG---AGTGGATCAACAGCCATTTCCACT 57
Db	235	ATGCTGAAATTCGAAGAGACTGCGAAGTGTGTGGCAGAAGACCCACAGCTCATCCCATG 294
Qy	58	TATCCAAAGACCTTGATTCGAAGAGATAGTCTTCAACAAAACCTGCGAGTGAAGT 117
Db	295	TACCCGACCCCTTGATTCGAAGAGATATGTTCTTCAGCAGAACTTGGCAGTGAAGT 354
Qy	118	TTTGGAACTGCTATCTGTGTTTCAGCAAGAAAGCCAAACAGGAGAGGAATTAAGGTA 177
Db	355	TTTGGAACTGCTATCTGTGTTTCAGCAAGAAAGCCAAACCTGAGAGGAGCTAAGGTA 414
Qy	178	CTTAAGGAATATCTGTGTGAGAACTTAATCCAAATGAACTGTACAGCCCAATTTGGA 237
Db	415	CTGAAGGAATCTCTGTGTGCGAATTAATCCAAATGAGACCGTGCAGGCCAATGGA 474
Qy	238	GCCCAACTCTCTCCAAAGCTGGACCCAGCCATTTGTCAGTTCATGCAAGTTTGTG 297
Db	475	GCTCAGCTCTCTCCAGCTCCATCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 534
Qy	298	GAGCAAGATATTTCTGCATATACAGAGTACTGTGAGGCGCGAGATCTGAGCAATAA 357
Db	535	GAGCAGGAGACTTTTTCATATCACGGAATACTGTGAGGCGCGAGATCTGGAATAGA 594
Qy	358	ATTGAGGAATATAACCAAGCTGGAATAATCTTCCAGAAATCAATATATAGAAATGTTT 417
Db	595	ATCCAGGAATATAAGAGCTGGAAAGTCTTCTGTAATAATCAGATAGTAAATGTTT 554
Qy	418	ATCCAGCTGCTGTGGAGTGTACTACATCATGAGAGGAGATCTTATCATGAGACTTA 477
Db	655	ATCCAGCTGTTGCTGTGAGTGTACTACATCATGAGAGGAGATCTTATCATGAGACTTG 714
Qy	478	AAGTCAAGAAATGATTTCTGAAATAATCTCCTTAAATTTGAGATTTGAGTTTCT 537
Db	715	AAATCAAGAAATATATTTCTGAAATAATCTTACTCAAAAT----- 755
Qy	538	CGACTTCTAATGGATCCTGTGACCTGGCCACAACTTTAACTGGAACTCCCATATATG 597
Db	756	----- 755
Qy	598	AGTCTGAGGCTGTGAACACCAAGGCTATGACACAAAGTCGACATCTGCTCACTGGCA 657
Db	756	-----CGGTCACTGGCA 767
Qy	658	TGCATTTTGTATGATGTGCTGATGAATCATGCAATTCGCTGGCTCCCAATTTCTTATCC 717
Db	768	TGCATTTTATATGATGTGCTGCTGATCATGCAATTCGCTGGCTCCAGTTTCTTGTCT 827
Qy	718	ATTGTTTTAAATTTGTTGAAGGTGACACACCTTCTCTCCCTGAGAGATATCCAAAGAA 777
Db	828	GTGTTTTGATATTTGTTGAAGGTAAACACCTTCACTCCCGACAGATATCCACAGAA 887
Qy	778	CTAATGCCATCATGGAAGCATGTTCAACAGAACTCTTCAATTAAGACCATCTGCTATC 837
Db	888	CTAATGCCATCATGGAAGCATGTTGUAACAGAGTCTTCAATTAAGACCATCTGCTATC 947
Qy	838	GAATTTTAAATTTCCCTTACCTTGTATGAGAGCTACAGAACCTTAATGTGTAGATTTCA 897
Db	948	GACATTTTAAAGGCCCTTACATGGAAGAGCAGCTTCAGTCTCTGTGTGTAATACCCA 1007
Qy	898	GAAATGACTCTGGAAGCAAAATTTGGATGTGTCAGAGGAGCTGCTCATATAATTAAT 957
Db	1008	GAGATGACTCTGGAAGCAAAATCTAGTTTGTGTCAGAGGAGCTGCTCATACATTAAT 1067
Qy	958	GCCATGCAAAAAGGATCCACTGTCAGACTCTGAGGCACTGTGAGAGTACAGAAATG 1017
Db	1068	GCCGTGCAAAAAGGATCCACTGTCAGACTCTGAGGCACTGTGAGAGTACAGAAATG 1127
Qy	1018	ACGCCAAGAAAGGATCGGCTGAGGAGCTCCAGCGGCTGATGGAAGCCAGGAG 1077
Db	1128	ACTCCAGAGAACGATGTGTGCTGAGGAAGCTGCGAGCAGCCGATGAGAGCCAGAGG 1187
Qy	1078	CTGAAAAGATTTGGAGAAAATATGAAGAAATAGCAAAATGCAAAATGCAAGATTTGAGA 1137
Db	1188	CTGAAAAGATTTGCTGAGAAAATATGAAGAAATGCAAAAGGATGCGAGCCCTGAGA 1247
Qy	1138	TCTCGGAACCTTTCAGCAGCTGAGTGTGATGTACTCCATGAA 1179
Db	1248	TCCCGAATGTGGGTCCGTGCATGCCCATGTGCTCCATGAA 1289
RESULT 9		
BC009414		
LOCUS		
DEFINITION		674 bp mRNA linear HTC 19-NOV-2003
ACCESSION		BC009414
VERSION		BC009414.1
KEYWORDS		GI:14714382
SOURCE		HTC.
ORGANISM		Homo sapiens (human)
REFERENCE		1 (bases 1 to 674)
AUTHORS		Strausberg, R.L., Feingold, E.A., Grouse, L.H., Derge, J.G., Klausner, R.D., Collins, F.S., Wagner, L., Shenmen, C.M., Schuler, G.D., Altschul, S.F., Zeeberg, B., Buetow, K.H., Schaefer, C.F., Bhat, N.K., Hopkins, R.F., Jordan, H., Moore, T., Max, S.I., Wang, J., Hsieh, F., Diatchenko, L., Marusina, K., Farmer, A.A., Rubin, G.M., Hong, L., Stapleton, M., Soares, M.B., Bonaldo, M.F., Casavant, T.L., Scheetz, T.E., Brownstein, M.J., Usdin, T.B., Toshiyuki, S., Carninci, P., Prange, C., Raha, S.S., Loquellano, N.A., Peters, G.J., Abramson, R.D., Mullany, S.J., Bosak, S.A., McEwen, P.J., McKernan, K.J., Malek, J.A., Gunaratne, P.H., Richards, S., Worley, K.C., Hale, S., Garcia, A.M., Gay, L.J., Hulyk, S.W., Villalón, D.M., Sodergren, E.J., Lu, X., Gibbs, R.A., Pahey, J., Helton, E., Kettman, M., Madan, A., Rodriguez, S., Sanchez, A., Whiting, M., Madan, A., Young, A.C., Shevchenko, Y., Bouffard, G.G., Blakesley, R.W., Touchman, J.W., Green, E.D., Dickson, M.C., Rodriguez, A.C., Grimwood, J., Schmutz, J., Myers, R.M., Butterfield, Y.S., Krzywinski, M.I., Skalska, U., Smal, D.E., Schnerch, A., Schein, J.E., Jones, S.J. and Marra, M.A.
TITLE		Generation and initial analysis of more than 15,000 full-length human and mouse cDNA sequences
JOURNAL		Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)
MEDLINE		22388257
PUBMED		12477932
REFERENCE		2 (bases 1 to 674)
AUTHORS		Strausberg, R.
TITLE		Direct Submission
JOURNAL		Submitted (19-JUN-2001) National Institutes of Health, Mammalian Gene Collection (MGC), Cancer Genomics Office, National Cancer Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590, USA
REMARK		NIH-MGC Project URL: <a href="http://mgc.nci.nih.gov">http://mgc.nci.nih.gov</a>
COMMENT		Contact: MGC help desk Email: <a href="mailto:cgapbs-re@mail.nih.gov">cgapbs-re@mail.nih.gov</a> Tissue Procurement: ATCC CDNA Library Preparation: Rubin Laboratory CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL) BC sequencing by: Genome Sequence Centre, BC Cancer Agency, Vancouver, BC, Canada <a href="mailto:info@bcgsc.bc.ca">info@bcgsc.bc.ca</a> Steven Jones, Jennifer Asano, Ian Bosdet, Yaron Butterfield, Susanna Chan, Readman Chiu, Chris Fjell, Erin Garland, Ran Guin, Letticia Hsiao, Martin Krzywinski, Reta Kutsche, Oliver Lee, Soo Sen Lee, Victor Ling, Carrie Mathewson, Candice McLeavy, Steven Ness, Pawan Pandoh, Anna-Liisa Prabhu, Parvaneh Saeedi, Jacqueline Schein, Duane Smalhus, Michael Smith, Lorraine Spence, Jeff Stott, Michael Thorne, Miranada Tsai, Natasja van den Bosch, Jill Vardy, George Yang, Scott Zuyderduyn, Marco Marra.

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>  
Series: IRAL Plate: 23 Row: p Column: 9  
This clone has the following problem: retained intron.

## FEATURES

1..674  
Location/Qualifiers  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="IMAGE:3613037"  
/tissue\_type="uterus, endometrium adenocarcinoma"  
/clone\_lib="NIH MGC 44"  
/lab\_host="DH10B-R"  
/note="Vector: pOTB7"

## ORIGIN

Query Match 42.4%; Score 614; DB 11; Length 674;  
Best Local Similarity 98.1%; Pred. No. 8.4e-136;  
Matches 632; Conservative 0; Mismatches 10; Indels 2; Gaps 1;

```
QY 170 TAAAGTACTTAAGGAATATCTGTGGAGACTAAATCCCAATGAACTGTACAGGCCA 229
DB 1 TAAAGTACTTAAGGAATATCTGTGGAGACTAAATCCCAATGAACTGTACAGGCCA 60
QY 230 ATTTCGAGCCCACTCTCTCCAAAGCTGGACCCAGCCCACTTGTCAAGTTCCATGCAA 289
DB 61 ATTTCGAGCCCACTCTCTCCAAAGCTGGACCCAGCCCACTTGTCAAGTTCCATGCAA 120
QY 290 GTTTTGTGAGCAAGATAATTTCTGCATATACCGAGTACTGTGAGGCCGAGATCTGG 349
DB 121 GTTTTGTGAGCAAGATAATTTCTGCATATACCGAGTACTGTGAGGCCGAGATCTGG 180
QY 350 ACGATAAATTCAGGAATATAAACAAGCTGGAAAAATCTTCCAGAAAAATCAATAATAG 409
DB 181 ACGATAAATTCAGGAATATAAACAAGCTGGAAAAATCTTCCAGAAAAATCAATAATAG 240
QY 410 AATGTTTATCAGCTGTCTGGAGTTGACTACATCATGATGAGAGGATCTTCATC 469
DB 241 AATGTTTATCAGCTGTCTGGAGTTGACTACATCATGATGAGAGGATCTTCATC 300
QY 470 GAGACTTAAAGTCAAGATATTTCTGAAATAATCTCTTAAATTTGAGATTTTG 529
DB 301 GAGACTTAAAGTCAAGATATTTCTGAAATAATCTCTTAAATTTGAGATTTTG 360
QY 530 GAGTTTCTCGACTTCTAATGGGATCTGTGACCTGGCCACAACTTTAAGTGGAACTCCCC 589
DB 361 GAGTTTCTCGACTTCTAATGGGATCTGTGACCTGGCCACAACTTTAAGTGGAACTCCCC 420
QY 590 ATTATATGATCTGAGGCTCTGAACACCAAGGCTATGACACAAAGTCCGACATCTGGT 649
DB 421 ATTATATGATCTGAGGCTCTGAACACCAAGGCTATG--ACAAAGTCCGACATCTGGT 478
QY 650 CACTGGCATGCATTTTGTATGATGTGCTGCATGAATCATGATTCGGTGGCTCCAAAT 709
DB 479 CACTGGCATGCATTTTGTATGATGTGCTGCATGAATCATGATTCGGTGGCTCCAAAT 538
QY 710 TCTTATCCATTTGTTTAAAAATTTGTAAGGTGACACACCTTCTCTCCCTGAGATATC 769
DB 539 TCTTATCCATTTGTTTAAAAATTTGTAAGGTGACACACCTTCTCTCCCTGAGATATC 598
QY 770 CAAAGAACTAATGCGCATGCAAGCATGTTGAAACAAGAT 813
DB 599 CAAAGAACTAATGCGCATGCAAGAAAGGTATAGAAATAACAT 642
```

RESULT 10  
BE388672  
LOCUS 601286319F1 NIH\_MGC\_44 Homo sapiens cDNA clone IMAGE:3613037 5',  
DEFINITION mRNA sequence.  
ACCESSION BE388672  
VERSION BE388672.1 GI:9334037  
KEYWORDS EST.

## SOURCE

ORGANISM Homo sapiens (human)  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

## REFERENCE

1 (Bases 1 to 671)

## AUTHORS

NIH-MGC <http://mgc.nci.nih.gov/>.

## TITLE

National Institutes of Health, Mammalian Gene Collection (MGC)

## JOURNAL

Unpublished (1999)

## COMMENT

Contact: Robert Strausberg, Ph.D.  
Email: [cgapbs-remail.nih.gov](mailto:cgapbs-remail.nih.gov)  
Tissue Procurement: ATCC  
cDNA Library Preparation: Ling Hong/Rubin Laboratory  
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)  
DNA sequencing by: Incyte Genomics, Inc.  
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at:  
<http://image.llnl.gov>  
Plate: LLCM273 row: e column: 06  
High quality sequence stop: 602.

## FEATURES

Location/Qualifiers

1..671

source

/organism="Homo sapiens"

/mol\_type="mRNA"

/db\_xref="taxon:9606"

/clone="IMAGE:3613037"

/tissue\_type="endometrium, adenocarcinoma cell line"

/lab\_host="DH10B (phage-resistant)"

/clone\_lib="NIH\_MGC\_44"

/note="Organ: uterus; Vector: pOTB7; Site 1: XhoI; Site 2: EcoRI; cDNA made by oligo-dT priming. Directionally cloned into EcoRI/XhoI sites using the following 5' adaptor: GGCAAGAG(G). Library constructed by Ling Hong in the laboratory of Gerald M. Rubin (University of California, Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and Superscript II RT (Life Technologies)."

## ORIGIN

Query Match 39.4%; Score 570.2; DB 10; Length 671;  
Best Local Similarity 97.1%; Pred. No. 2.4e-125;  
Matches 612; Conservative 0; Mismatches 14; Indels 4; Gaps 3;

```
QY 189 ATCTGTTGGAGAACTAAATCCAAATGAACTGTACAGGCCCAATTTGGAAGCCCACTCT 248
DB 1 ATCTGTTGGAGAACTAAATCCAAATGAACTGTACAGGCCCAATTTGGAAGCCCACTCT 60
QY 249 CTCCAAGCTGACACACCCAGCCATTGTCAAGTTCCATGCCAAGTTTGTGAGCAAGATA 308
DB 61 CTCCAAGCTGACACACCCAGCCATTGTCAAGTTCCATGCCAAG-TTTGTGAGCAAGATA 119
QY 309 TTTCTGCATATACCGAGTACTGTGAGGCCCGAGATCTGGACGATAAATTCAGGAATA 368
DB 120 TTTCTGCATATACCGAGTACTGTGAGGCCCGAGATCTGGACGATAAATTCAGGAATA 179
QY 369 TAAACAAGCTGAAAAATCTTTCCAGAAAAATCAAAATAATAGAATGTTTATCCAGCTGCT 428
DB 180 TAAACAAGCTGAAAAATCTTTCCAGAAAAATCAAAATAATAGAATGTTTATCCAGCTGCT 239
QY 429 GCTGGAGTTGACTACATGATGATGAGAGGATATCTTCATCGACACTTAAGTCAAGAA 488
DB 240 GCTGGAGTTGACTACATGATGATGAGAGGATATCTTCATCGACACTTAAGTCAAGAA 299
QY 489 TGTATTTCTGAAAAATAATCTCTTAAATTTGGAGATTTTGGAGTTTCTCGACTCTAAT 548
DB 300 TGTATTTCTGAAAAATAATCTCTTAAATTTGGAGATTTTGGAGTTTCTCGACTCTAAT 359
QY 549 GGGATCTCTGTGACCTGGCCCAACTTTTAAGTGGAACTCCCATTAATATGAGTCTCTGAGGC 608
DB 360 GGGATCTCTGTGACCTGGCCCAACTTTTAAGTGGAACTCCCATTAATATGAGTCTCTGAGGC 419
QY 609 TCTGAACACCAAGCTATGACACAAAGTGGACATCTGGTCACTGGCATGCACTTTTGTA 668
DB 420 TCTGAACACCAAGGCTATG--ACAAAGTGGACATCTGGTCACTGGCATGCA--TTTGTA 476
```



```

QY 669 TGAGATGCTGCATGATCATGATTCGCTGCTGCTCAATTCCTTATCCATTGTTTAA 728
|
|
|
Db 477 TGAGATGCTGCATGATCATGATTCGCTGCTGCTCAATTCCTTATCCATTGTTTAA 536
|
|
|
QY 729 AATGTTTGAAGGTGACACACCTTCTCTCCCTGAGAGATATCCAAAAGAACTAAATGCCAT 788
|
|
|
Db 537 AATCGTTGAAGGTGACACACCTTCTCTCCCTGAGAGATATCCAAAAGAACTAAATGCCAT 596
|
|
|
QY 789 CATGGAAGCATGTTGACCAAGATCCTTC 818
|
|
|
Db 597 CATGGAAGCATGTTGACCAAGATGATGTC 626
|
|
|
RESULT 11
BG752438 659 bp mRNA linear EST 15-MAY-2001
LOCUS 602730927F1 NIH_MGC_43 Homo sapiens cDNA clone IMAGE:4874479 5',
DEFINITION mRNA sequence.
ACCESSION BG752438
VERSION BG752438.1 GI:14063091
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE NIH-MGC http://mgs.nci.nih.gov/.
AUTHORS National Institutes of Health, Mammalian Gene Collection (MGC)
TITLE Unpublished (1999)
JOURNAL Contact: Robert Strausberg, Ph.D.
COMMENT Email: cgapbs-remail.nih.gov
Tissue Procurement: ATCC
cDNA Library Preparation: Ling Hong/Rubin Laboratory
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLC1754 row: e column: 08
High quality sequence stop: 657.
FEATURES
source
1..659
Location/Qualifiers
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:4874479"
/tissue_type="normal pigmented retinal epithelium"
/lab_host="DH10B (phage-resistant)"
/clone_lib="NIH_MGC_43"
/note="Organ: eye; Vector: pOTB7; Site 1: XhoI; Site 2:
EcoRI; cDNA made by oligo-dT priming. Directionally
cloned into EcoRI/XhoI sites using the following 5'
adaptor: GGACAG(G). Library constructed by Ling Hong
in the laboratory of Gerald M. Rubin (University of
California, Berkeley) using ZAP-cDNA synthesis kit
(Stratagene) and Superscript II RT (Life Technologies).
Note: this is a NIH_MGC Library. |"
ORIGIN
Query Match 39.0%; Score 564.6; DB 12; Length 659;
Best Local Similarity 97.1%; Pred. No. 5.3e-124;
Matches 640; Conservative 0; Mismatches 9; Indels 10; Gaps 6;
QY 612 GAAACACCAAGGTATGACACAAAGTCGGACATCTGGTCACTGGCATGCAATTTGTATGA 671
|
|
|
Db 1 GAAACACCAAGGTATGACACAAAGTCGGACATCTGGTCACTGGCATGCAATTTGTATGA 60
|
|
|
QY 672 GATGTGCTCAGCAATCATGCAATCGCTGGCTCCCAATTCCTATCCATTGTTTAAAT 731
|
|
|
Db 61 GATGTGCTCAGCAATCATGCAATCGCTGGCTCCCAATTCCTATCCATTGTTTAAAT 120
|
|
|
QY 732 TGTGGAAGTGACACACCTTCTCTCCCTGAGAGATATCCAAAAGAACTAAATGCCATCAT 791
|
|
|

```

```

Db 121 TGTGAAGGTGACACACCTTCTCTCCCTGAGAGATATCCAAAAGAACTAAATGCCATCAT 180
|
|
|
QY 792 GGAAGCATGTTGAACCAAGAAATCCTTCAATTAAGAACCATCTGCTATCGAAATTTTAAAAAT 851
|
|
|
Db 181 GGAAGCATGTTGAACCAAGAAATCCTTCAATTAAGAACCATCTGCTATCGAAATTTTAAAAAT 240
|
|
|
QY 852 CCCTTACCTTGTATGAGCAGCTACAGAACCTTAATGTGTA-GATATTAGAAATGACTCTGG 910
|
|
|
Db 241 CCCTTACCTTGTATGAGCAGCTACAGAACCTTAATGTGTACGATATTTCAGAAATGACTCTGG 300
|
|
|
QY 911 AAGACAAAATTTGGATTGTGACAGAGGAGCTGCTCATATAATTAATGCCATGCAAAAAA 970
|
|
|
Db 301 AAGACAAAATTTGGATTGTGACAGAGGAGCTGCTCATATAATTAATGCCATGCAAAAAA 360
|
|
|
QY 971 GGATCCACCTGACAGCTCTGAGGGCACTGTCAAGATPACAGAAAATGACCCCAAGAGAAA 1030
|
|
|
Db 361 GGATCCACCTGACAGCTCTGAGGGCACTGTCAAGATPACAGAAAATGACCCCAAGAGAAA 420
|
|
|
QY 1031 GGATCGGCTGAGGAAGCTCCAGCGGCTCATGAGAAAGCCAGAGAGCTGAA-AAAGATT 1089
|
|
|
Db 421 GGATCGGCTGAGGAAGCTCCAGCGGCTCATGAGAAAGCCAGAGAGCTGAA-AAAGATT 480
|
|
|
QY 1090 GTGGAAGAAAATATGAAGAAAATAGCAAAACGAATGCAAGAA-TTGAGATCTCGGAATT 1148
|
|
|
Db 481 GTGGAAGAAAATATGAAGAAAATAGCAAAACGAATGCAAGAACTTGAGATCTCGGAATT 540
|
|
|
QY 1149 TCAGCAGCTGAGTGTTCATGCTACTCCATGAAAAACACATTT---AAAGGATGG--AA 1203
|
|
|
Db 541 TCAGCAGCTGAGTGTTCATGCTACTCCATGAAAAACACATTTTACAAAGGAATGACAG 600
|
|
|
QY 1204 GAAAAGGAGGAGCAACCTGGA--GGGAAGACTTTCTTGTTCACCCAGGACGAGATGAA 1260
|
|
|
Db 601 AAAAGGAGGAGCAACCTGGAAGGAGAGACTTTTCTTGTTCACCCAGGACGAGATGAA 659
|
|
|
RESULT 12
BI544364 767 bp mRNA linear EST 05-SEP-2001
LOCUS 603241760F1 NIH_MGC_95 Homo sapiens cDNA clone IMAGE:5284263 5',
DEFINITION mRNA sequence.
ACCESSION BI544364
VERSION BI544364.1 GI:15431676
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE NIH-MGC http://mgs.nci.nih.gov/.
AUTHORS National Institutes of Health, Mammalian Gene Collection (MGC)
TITLE Unpublished (1999)
JOURNAL Contact: Robert Strausberg, Ph.D.
COMMENT Email: cgapbs-remail.nih.gov
Tissue Procurement: Miklos Palkovits, M.D., Ph.D.
cDNA Library Preparation: Michael J. Brownstein (NHGRI), Shiraki
Toshiyuki and Piero Carninci (RIKEN)
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLAM1718 row: g column: 16
High quality sequence stop: 728.
FEATURES
source
1..767
Location/Qualifiers
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:5284263"
/tissue_type="hippocampus"
/lab_host="DH10B"
/clone_lib="NIH_MGC_95"
/note="Organ: brain; Vector: pBluescriptR (modified)

```



pBluescript KS+); Site 1: BamHI; Site 2: SalI-XhoI (gtcag); Oligo-dr primed using primer 5'-TTTTTTTTTTTTTTN-3', size-selected for average insert size 2.5 kb and normalized to ROT 5. This is a primary library enriched for full-length clones and constructed using the Cap-trapper method (Carninci, in preparation). Library constructed by M. Brownstein (NIH/NHGRI, National Institutes of Health). Note: this is a NIH\_MGC Library."

```

ORIGIN
Query Match      35.2%; Score 510.6; DB 12; Length 767;
Best Local Similarity 98.6%; Pred. No. 4.2e-111;
Matches 547; Conservative 0; Mismatches 4; Indels 4; Gaps 3;

QY 1 ATGCTGAATTCGAAGAGCGAGCTTAAGTGTGTGAGTGATCAACAGCCATTTCCACTTAT 60
DB 216 ATGCTGAATTCGAAGAGCGAGCTTAAGTGTGTGAGTGATCAACAGCCATTTCCACTTAT 275

QY 61 CCAAGAGCTTGTATTCGAAGAGATACGTGTTCAACAAAACCTTGGCAGTGGAGTTT 120
DB 276 CCAAGAGCTTGTATTCGAAGAGATACGTGTTCAACAAAACCTTGGCAGTGGAGTTT 335

QY 121 GGAATGTCTATCTGTTTTCAGACAAGAACGCAACGAGGAGGAATTAAGTACTT 180
DB 336 GGAATGTCTATCTGTTTTCAGACAAGAACGCAACGAGGAGGAATTAAGTACTT 395

QY 181 AAGGAATATCTGTTGGAGAACTAATCCAAATGAACTGTACAGGCCAATTTGAGGCC 240
DB 396 AAGGAATATCTGTTGGAGAACTAATCCAAATGAACTGTACAGGCCAATTTGAGGCC 455

QY 241 CAATCTCTCTCAAGCTGGACCCAGCCATTTGTCAGAGTTCCATGCAAGTTTGTGGAG 300
DB 456 CAATCTCTCTCAAGCTGGACCCAGCCATTTGTCAGAGTTCCATGCAAGTTTGTGGAG 515

QY 301 CAAGTAATTTCTGATATACGAGAGTACTGTGAGGCCCGAGATCTGAGAGTAATTT 360
DB 516 CAAGTAATTTCTGATATACGAGAGTACTGTGAGGCCCGAGATCTGAGAGTAATTT 575

QY 361 CAGGAATATAACAAGCTGGAATAATCTTTCAGAAAATCAATAATAGATGGTTTATC 420
DB 576 CAGGAATATAACAAGCTGGAATAATCTTTCAGAAAATCAATAATAGATGGTTTATC 635

QY 421 CAGCTGCTGCTGGAGTTGAC-TACATGATGAGAGGAGGATACCTTCATCGAGACTTAA 479
DB 636 CAGCTGCTGCTGGAGTTGACTTACATGATGAGAGGAGGATACCTTCATCGAGACTT-AA 694

QY 480 GTCAAGATGTATTTCTGAAAATAATCTCTTAAATTTGGAGATTTTTCG 539
DB 695 GTCAAGATGTATTTCTGAAAATAATCTCTTAAATTTGGAGATTTTTCG 752

QY 540 ACTTCTAATGGGATC 554
DB 753 ACTTCTAATGGGATC 767

```

```

RESULT 13
BU930538
LOCUS
DEFINITION BU930538 867 bp mRNA linear EST 18-OCT-2002
AGENCY 10424759 NIH_MGC_83 Homo sapiens cDNA clone IMAGE:6668583
5' mRNA sequence.
ACCESSION BU930538
VERSION BU930538.1 GI:24119357
KEYWORDS EST.
SOURCE Homo sapiens
ORGANISM Homo sapiens (human)
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 867)
NIH-MGC http://mgi.nci.nih.gov/
AUTHORS National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.

```

Email: cgabbs@mail.nih.gov  
Tissue Procurement: CLONTECH  
CDNA Library Preparation: CLONTECH Laboratories, Inc.  
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)  
DNA Sequencing by: Agencourt Bioscience Corporation  
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at:  
http://image.llnl.gov  
Plate: LLCM2942 row: 0 column: 15  
High quality sequence stop: 424.

#### FEATURES

Location/Qualifiers  
1..867  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="IMAGE:6668583"  
/lab\_host="NIH\_MGC\_83"  
/clone\_lib="NIH\_MGC\_83"  
/notes="Organ: prostate; Vector: pDNR-LIB (Clontech);  
Site 1: SfII (ggccgctggcc); Site 2: SfII  
(ggccattatggcc); 5' and 3' adaptors were used in cloning  
as follows: 5' adaptor sequence: 5'-CACGGCCATTATGGCC-3'  
and 3' adaptor sequence:  
5'-ATTCTAGAGCGGCGGCGGCATG-dT(30)BN-3' (where B = A,  
C, or G and N = A, C, G, or T). Average insert size 1.4  
Kb (range 0.5-4.0 Kb). 14/15 colonies contained inserts  
by PCR. This library was enriched for full-length clones  
and was constructed by Clontech Laboratories (Palo Alto,  
CA)."

#### ORIGIN

```

Query Match      35.2%; Score 509.6; DB 13; Length 867;
Best Local Similarity 98.0%; Pred. No. 7.4e-111;
Matches 537; Conservative 0; Mismatches 9; Indels 2; Gaps 2;

QY 331 TGTGAGGCCGAGATCTGGACGATATAAATTCAGGAATATAACAAGCTGGAAAAATCTTT 390
DB 38 TTTGAGGCCGAGATCTGGACGATATAAATTCAGGAATATAACAAGCTGGAAAAATCTTT 97

QY 391 CCAGAAATCAATAATAGAAATGTTTATCCAGCTGCTGCGAGTTGACTACATGCAT 450
DB 98 CCAGAAATCAATAATAGAAATGTTTATCCAGCTGCTGCGAGTTGACTACATGCAT 157

QY 451 GAGAGGAGGATACCTTCATCGAGACTTAAAGTCAAGAATGTAATTTCTGAAAAATAATCTC 510
DB 158 GAGAGGAGGATACCTTCATCGAGACTTAAAGTCAAGAATGTAATTTCTGAAAAATAATCTC 217

QY 511 CTTAAAAATTGGAGATTTTGGAGTTTTCGACTTCTAATGGGATCCTGTGACCTGGCCACA 570
DB 218 CTTAAAAATTGGAGATTTTGGAGTTTTCGACTTCTAATGGGATCCTGTGACCTGGCCACA 277

QY 571 ACTTAACTGGAGTCCGCCATTTATGAGTCTCTGAGGCTCTGAAACACCAAGGCTATGAC 630
DB 278 ACTTAACTGGAGTCCGCCATTTATGAGTCTCTGAGGCTCTGAAACACCAAGGCTATGAC 337

QY 631 ACAAGTCCGACATCTGGTCACTGGCATGCAATTTTGTATGAGATGTCTGCATGAATCAT 690
DB 338 ACAAGTCCGACATCTGGTCACTGGCATGCAATTTTGTATGAGATGTCTGCATGAATCAT 397

QY 691 GCATTCGCTGGCTCCAATTTCTTATCCATTTTAAATTTTAAATTTGTAAGCTGACACCT 750
DB 398 GCATTCGCTGGCTCCAATTTCTTATCCATTTTAAATTTTAAATTTGTAAGCTGACACCT 457

QY 751 TCTCTCCCTGAGATATCCAAAAGAACTAAATGCCATCATGGAAGCAATGTTTGAACAAG 810
DB 458 TCTCTCCCTGAGATATCCAAAAGAACTAAATGCCATCATGGAAGCAATGTTTGAACAAG 517

QY 811 AATCC-TTCATTAAAGACATCTGTATCGAAATTTTAAAT-CCCTTACCTTATGAC 868
DB 518 AATCC-TTCATTAAAGACATCTGTATCGAAATTTTAAAT-CCCTTACCTTATGAC 577

QY 869 AGCTACAG 876
|||||

```

Db 578 AGTACAG 585

RESULT 14  
CD106568  
LOCUS  
DEFINITION  
ACCESSION  
VERSION  
KEYWORDS  
SOURCE  
ORGANISM

CD106568 845 bp mRNA linear EST 15-MAY-2003  
AGENCOURT\_13979292 NIH MGC 179 Homo sapiens cDNA clone  
IMAGE:30368885 5', mRNA sequence.  
CD106568  
CD106568.1 GI:30759742  
EST.  
Homo sapiens (human)  
Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
1 (bases 1 to 845)  
NIH-MGC <http://mgi.nci.nih.gov/>  
National Institutes of Health, Mammalian Gene Collection (MGC)  
Unpublished (1999)  
Contact: Robert Strausberg, Ph.D.  
Email: [cgabs@mail.nih.gov](mailto:cgabs@mail.nih.gov)  
Tissue Procurement: Dr. Michael Brownstein  
cDNA Library Preparation: Invitrogen Corp  
DNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)  
DNA Sequencing by: Agencourt Bioscience Corporation  
Clone distribution: MGC clone distribution information can be  
found through the I.M.A.G.E. Consortium/LLNL at:  
<http://image.llnl.gov>  
Plate: NDAM432 row: 9 column: 06  
High quality sequence stop: 719.  
Location/Qualifiers  
1..845  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="IMAGE:30368885"  
/issue\_type="Pituitary"  
/lab\_host="PH108-Ton A ( T1 and T5 phage resistances)"  
/clone\_lib="NIH\_MGC\_179"  
/note="Organ: Brain; Vector: pCMV-sport6.1; Site 1: EcoRV  
(destroyed); Site 2: NotI; Library is oligo-dT primed and  
directionally cloned (EcoRV site is destroyed upon  
cloning). Average insert size 1.1 kb. Library was  
constructed by (Invitrogen). Note: this is a NIH\_MGC  
Library."

FEATURES  
source

ORIGIN

Query Match 34.8%; Score 504; DB 14; Length 845;  
Best Local Similarity 100.0%; Pred. No. 1.6e-109;  
Matches 504; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 896 CAGAATGACTCTGGAAGACAAAATTTGGATTGTCAGAGGAGGCTGCTCATATAATTA 955  
DB 1 CAGAATGACTCTGGAAGACAAAATTTGGATTGTCAGAGGAGGCTGCTCATATAATTA 60  
QY 956 ATGCCATGCAAAAAGGATCCACCTGCGAGACTCTGAGGCGCACTGTGAGAACTACAGAAA 1015  
DB 61 ATGCCATGCAAAAAGGATCCACCTGCGAGACTCTGAGGCGCACTGTGAGAACTACAGAAA 120  
QY 1016 TCAGCCCAAGAGAGGATGCGGCTGAGAGAGCTCCAGCGGCTGTAGAGAAACCCAGGA 1075  
DB 121 TCAGCCCAAGAGAGGATGCGGCTGAGAGAGCTCCAGCGGCTGTAGAGAAACCCAGGA 180  
QY 1076 AGCTGAAAAGATTTGGAAGAAAATATGAAGAAAATAGCAAAACGAAATGGA 1135  
DB 181 AGCTGAAAAGATTTGGAAGAAAATATGAAGAAAATAGCAAAACGAAATGGA 240  
QY 1136 GATCTCGGAACCTTCAGGAGCTGAGTGTGTGATCTTCATGAAAAACACATTTAAAG 1195  
DB 241 GATCTCGGAACCTTCAGGAGCTGAGTGTGTGATCTTCATGAAAAACACATTTAAAG 300  
QY 1196 GAATGGAAGAAAAGGAGGAGCAACTGAGGGAAGACTTTCTTGTTCACCCAGGACGAGG 1255

Db 301 GAATGGAAGAAAAGGAGGAGCAACTGAGGGAAGACTTTCTTGTTCACCCAGGACGAGG 360  
QY 1256 ATGAAGAGAGGTGCAAGGCGAGGGAAGAGGAAATCTGATGAACCAACTTTAGAGAACTGCG 1315  
DB 361 ATGAAGAGAGGTGCAAGGCGAGGGAAGAGGAAATCTGATGAACCAACTTTAGAGAACTGCG 420  
QY 1316 CTGAGTCTCAGCCCTATTCTCTTCATGAGCTCCACGAACTTGAATCAATTTGTAGAGATG 1375  
DB 421 CTGAGTCTCAGCCCTATTCTCTTCATGAGCTCCACGAACTTGAATCAATTTGTAGAGATG 480  
QY 1376 CCACATCTGACCTTTGGATACCATG 1399  
DB 481 CCACATCTGACCTTTGGATACCATG 504

RESULT 15  
BB552498  
LOCUS  
DEFINITION  
ACCESSION  
VERSION  
KEYWORDS  
SOURCE  
ORGANISM

BB552498 635 bp mRNA linear EST 26-OCT-2001  
BB552498 RIKEN full-length enriched, 2 days pregnant adult female  
oviduct Mus musculus cDNA clone E230038G21 3', mRNA sequence.  
BB552498  
BB552498.2 GI:16448169  
EST.  
Mus musculus (house mouse)  
Mus musculus  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
1 (bases 1 to 635)  
Arakawa, T., Carninci, P., Fukuda, S., Furuno, M., Hanagaki, T.,  
Hara, A., Hiramoto, K., Hori, F., Ishii, Y., Ito, M., Kawai, J.,  
Konno, H., Kouda, M., Koya, S., Matsuyama, T., Miyazaki, A., Nomura, K.,  
Ono, M., Okazaki, Y., Okido, T., Saito, R., Sakai, C., Sakai, K.,  
Sano, H., Sasaki, D., Shibata, K., Shinagawa, A., Shiraki, T.,  
Sogabe, Y., Suzuki, H., Tagami, M., Tagawa, A., Takahashi, F.,  
Takeda, Y., Tanaka, T., Toya, T., Muramatsu, M. and Hayashizaki, Y.  
RIKEN Mouse ESTs (Arakawa, T., et al. 2001)  
Unpublished (2001)  
On Aug 1, 2000 this sequence version replaced gi:9638864.  
Contact: Yoshihide Hayashizaki  
Laboratory for Genome Exploration Research Group, RIKEN Genomic  
Sciences Center (GSC), Yokohama Institute  
The Institute of Physical and Chemical Research (RIKEN)  
1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan  
Tel: 81-45-503-9222  
Fax: 81-45-503-9216  
Email: [genome-res@gsr.riken.go.jp](mailto:genome-res@gsr.riken.go.jp),  
[URL:http://genome.gsc.riken.go.jp/](mailto:URL:http://genome.gsc.riken.go.jp/)  
Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K.,  
Itoh, M., Konno, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.  
Normalization and subtraction of cap-trapper-selected cDNAs to  
prepare full-length cDNA libraries for rapid discovery of new  
genes. Genome Res. 10 (10), 1617-1630 (2000)  
wagi, K., Fujiwaka, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E.,  
Watahiki, M., Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T.,  
Matsuura, S., Kawai, J., Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A.  
and Hayashizaki, Y.  
RIKEN integrated sequence analysis (RISA) system--384-format  
sequencing pipeline with 384 multicapillary sequencer. Genome Res.  
10 (11), 1757-1771 (2000)  
Konno, H., Fukunishi, Y., Shibata, K., Itoh, M., Carninci, P.,  
Sugahara, Y. and Hayashizaki, Y.  
Computer-based methods for the mouse full-length cDNA  
encyclopedia: real-time sequence clustering for construction of a  
non-redundant cDNA library. Genome Res. 11 (2), 281-289 (2001)  
Kondo, S., Shinagawa, A., Saito, T., Kiyosawa, H., Yamakawa, I.,  
Aizawa, K., Fukuda, S., Hara, A., Itoh, M., Kawai, J., Shibata, K. and  
Hayashizaki, Y.  
Computational Analysis of Full-length Mouse cDNAs Compared with  
Human Genome Sequences Mamm. Genome. 12, 673-677 (2001)  
Please visit our web site (<http://genome.gsc.riken.go.jp/>) for  
further details.  
cDNA library was prepared and sequenced in Mouse Genome  
Encyclopedia Project of Genome Exploration Research Group in Riken

Genomic Sciences Center and Genome Science Laboratory in RIKEN.  
Division of Experimental Animal Research in Riken contributed to  
prepare mouse tissues.

FEATURES

source  
1..635  
/organism="Mus musculus"  
/mol\_type="mRNA"  
/db\_xref="taxon:10090"  
/clone="E230038G21"  
/sex="female"  
/tissue\_type="oviduct"  
/dev\_stage="2 days pregnant adult"  
/lab\_host="DH10B"  
/clone\_lib="RIKEN full-length enriched, 2 days pregnant  
adult female oviduct"  
/note="Site 1: SalI; Site 2: BamHI; cDNA library was  
prepared and sequenced in Mouse Genome Encyclopedia  
Project of Genome Exploration Research Group in Riken  
Genomic Sciences Center and Genome Science Laboratory in  
RIKEN. Division of Experimental Animal Research in Riken  
contributed to prepare mouse tissues. 1st strand cDNA was  
primed with a primer [5',  
GAGAGAGAGCGCGCAACTCGAGTTTCTTTTCTTTT 3'], cDNA was  
prepared by using trihalose thermo-activated reverse  
transcriptase and subsequently enriched for full-length by  
cap-trapper. Second strand cDNA was prepared with the  
primer adapter of sequence [5',  
GAGAGAGATCTCGAGTTAATAATTAATCCGCCCC 3']. cDNA  
was cleaved with BamHI and XhoI. Vector: a modified  
pbluescript KS(+) after Dulk excision from lambda FLC I."

ORIGIN

Query Match 32.1%; Score 465.2; DB 10; Length 635;  
Best Local Similarity 83.5%; Pred. No. 2.9e-100;  
Matches 527; Conservative 0; Mismatches 104; Indels 0; Gaps 0;  
QY 460 ATACTTCATCGAGACTTAAAGTCAAGATGATTTCTGAAAAATAATCTCCTTAATAATT 519  
Db 2 ATACTTCATCGAGACTTAAATCAAGATNATATTTCTGAAAAATAATCTACTCAAAATC 61  
QY 520 GGAGATTTTGGAGTTTCTCGACTTCTAATGGGATCTGTGACCTGGCCACAACCTTTAACT 579  
Db 62 GGGGATTTTGGAGTTTCTGGCTGTCTAATGGGTTCATGTGAGCTGGCTACACTCTAACC 121  
QY 580 GGAATCTCCCATATATAGTCTGAGGCTCTGAACAACAAGGGCTATGACACAAAGTCG 639  
Db 122 GGGACCCCCATATATAGTCCGAGGCCCTGAAGCACAAGGGCTATGATGCCAAGTCT 181  
QY 640 GACATCTGCTCACTGGCATGCAATTTGTATGAGTGTGCTGATGAATCATGCTTCGCT 699  
Db 182 GACATCTGCTCACTGGCATGCAATTTGTATGAGTGTGCTGATGAATCATGCTTCGCT 241  
QY 700 GGCTCCAAATTTCTATCCATTTGTTTAAATAATTTGAAGGTGACACACCTTCTCTCCCT 759  
Db 242 GGCTCCAGTTTCTTGCTGTGGTTTGAATATTTGTAAGGTAAACACCTTCACTCCCG 301  
QY 760 GAGAGATATCCAAAGAACTAATGCAATCGCAAGCATGTTGAACAAGATCCTTCA 819  
Db 302 GACAGATATCCAGAGAACTAAACCCATCATGGAACGCATGTTGAACAAGATCCTTCA 361  
QY 820 TTAAGACCATCTGCTATCGAAATTTTAAATAATCCCTTACCTTGATGAGCAGCTACAGAC 879  
Db 362 TTGAGACCGTCCGCTCGAGACATTTTAAAGCCCTTACATGAGAGCAGCTTCAGCTC 421  
QY 880 CTAATGTGTAGATATTAGAAATGACTCTGGAAGACAAAATTTGGATTGTGAGAGGAG 939  
Db 422 CTGATGTGTAAATACCAGAGATGACACTGGAAGACAAAGAACTCAGTTTTGTGAGAGGAG 481  
QY 940 GCTGCTCATATAATTAATGCCATGCAAAAAGGATCCACCTGCAGACTCTGAGGGCACTG 999  
Db 482 GCTGCTCATACGATTAATGCCGTGCAAAAAGAACTTCACTGCGACTCTGCAAGCCTTG 541  
QY 1000 TCAGAGTACAGAAATGACCCCAAGAGAAAGGATCGGCTGAGGAAGCTCCAGCGGCT 1059

Db 542 TCTGACACGCAGAAAAACGACTCCAGAGAACGATGTGGTGTGAGGAAGCTGCAGGAGCC 601  
QY 1060 GATGAGAAAGCCAGGAAGCTGAAAAAGATTG 1090  
Db 602 GATGAGAGAGCCAGGAGGCTGAAAAAGATTG 632

Search completed: October 3, 2004, 00:56:25  
Job time : 4050 secs

**This Page Blank (uspto)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: October 2, 2004, 23:49:06 ; Search time 740 Seconds  
(without alignments)  
9926.760 Million cell updates/sec

Title: US-10-803-278-3

Perfect score: 1449

Sequence: 1 atgttgaaattccaagagggc.....actggaaaaatgaaaaataa 1449

Scoring table: IDENTITY\_NUC

Gapop 10\_0 , Gapext 1.0

Searched: 3340653 seqs, 2534783454 residues

Total number of hits satisfying chosen parameters: 6691306

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications NA:\*

- 1: /cgn2\_6/ptodata/1/pubpna/US07\_PUBCOMB.seq\*
- 2: /cgn2\_6/ptodata/1/pubpna/PCT\_NEW\_PUB.seq\*
- 3: /cgn2\_6/ptodata/1/pubpna/US05\_NEW\_PUB.seq\*
- 4: /cgn2\_6/ptodata/1/pubpna/US06\_PUBCOMB.seq\*
- 5: /cgn2\_6/ptodata/1/pubpna/US07\_NEW\_PUB.seq\*
- 6: /cgn2\_6/ptodata/1/pubpna/PCTUS\_PUBCOMB.seq\*
- 7: /cgn2\_6/ptodata/1/pubpna/US08\_NEW\_PUB.seq\*
- 8: /cgn2\_6/ptodata/1/pubpna/US08\_PUBCOMB.seq\*
- 9: /cgn2\_6/ptodata/1/pubpna/US09A\_PUBCOMB.seq\*
- 10: /cgn2\_6/ptodata/1/pubpna/US09B\_PUBCOMB.seq\*
- 11: /cgn2\_6/ptodata/1/pubpna/US09C\_PUBCOMB.seq\*
- 12: /cgn2\_6/ptodata/1/pubpna/US09\_NEW\_PUB.seq\*
- 13: /cgn2\_6/ptodata/1/pubpna/US09\_NEW\_PUB.seq\*
- 14: /cgn2\_6/ptodata/1/pubpna/US10A\_PUBCOMB.seq\*
- 15: /cgn2\_6/ptodata/1/pubpna/US10B\_PUBCOMB.seq\*
- 16: /cgn2\_6/ptodata/1/pubpna/US10C\_PUBCOMB.seq\*
- 17: /cgn2\_6/ptodata/1/pubpna/US10\_NEW\_PUB.seq\*
- 18: /cgn2\_6/ptodata/1/pubpna/US60\_NEW\_PUB.seq\*
- 19: /cgn2\_6/ptodata/1/pubpna/US60\_PUBCOMB.seq\*

pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1449	100.0	1448	17	US-10-803-278-3
2	1399	96.5	1938	17	Sequence 3, Appli
3	1399	96.5	2552	13	Sequence 1, Appli
4	1399	96.5	2661	13	Sequence 25, Appli
5	1399	96.5	2809	13	Sequence 2275, A
6	1399	96.5	2847	17	Sequence 18976, A
7	1397.4	96.4	2869	9	Sequence 1, Appli
8	1397.4	96.4	2869	17	Sequence 1, Appli
9	1385.4	95.6	2765	13	Sequence 50, Appli
10	759	52.4	2484	13	Sequence 172, A
11	669.6	46.2	2538	10	Sequence 511, App
12	669.6	46.2	2538	15	Sequence 511, App
13	415.8	28.7	451	11	Sequence 60, Appli
14	395.4	27.3	430	9	Sequence 241, App

15	395.4	27.3	430	11	US-09-764-875-376	Sequence 376, App
16	199	13.7	1438	10	US-09-764-891-2042	Sequence 2042, App
17	172.2	11.9	326014	9	US-09-731-231A-3	Sequence 3, Appli
18	172.2	11.9	326014	17	US-10-751-985-3	Sequence 3, Appli
19	156	10.8	2022	10	US-09-764-891-7352	Sequence 7352, App
20	156	10.8	2961	10	US-09-764-891-7353	Sequence 7353, App
21	149.6	10.3	3696	17	US-10-620-052A-31	Sequence 31, Appli
22	149.4	10.3	1781	13	US-10-114-270-181	Sequence 181, App
23	149.4	10.3	2257	13	US-10-114-270-179	Sequence 179, App
24	148.2	10.2	1588	13	US-10-114-270-183	Sequence 183, App
25	148	10.2	3544	13	US-10-425-114-26260	Sequence 26260, A
26	148	10.2	3544	13	US-10-425-114-26275	Sequence 26275, A
27	148	10.2	3698	13	US-10-343-887-880	Sequence 880, App
28	148	10.2	3698	13	US-10-172-118-880	Sequence 880, App
29	147.8	10.2	2110	13	US-10-415-011-29	Sequence 29, Appli
30	143.2	9.9	3643	9	US-09-783-320-3	Sequence 3, Appli
31	143.2	9.9	5426	9	US-09-783-320-50	Sequence 50, Appli
32	143.2	9.9	5448	15	US-10-037-270-246	Sequence 246, App
33	143.2	9.9	5448	16	US-10-117-722-246	Sequence 246, App
34	143.2	9.9	5532	15	US-10-037-270-245	Sequence 245, App
35	143.2	9.9	5532	16	US-10-117-722-245	Sequence 245, App
36	139.4	9.6	418	10	US-09-764-891-1038	Sequence 1038, App
37	133.2	9.2	1587	13	US-10-114-270-187	Sequence 187, App
38	131.6	9.1	1965	9	US-09-940-921B-3	Sequence 3, Appli
39	131.6	9.1	2052	9	US-09-940-921B-1	Sequence 1, Appli
40	131.6	9.1	2240	9	US-09-940-921B-5	Sequence 5, Appli
41	131.6	9.1	2370	15	US-10-243-735-1	Sequence 1, Appli
42	131.6	9.1	2370	17	US-10-730-010-1	Sequence 1, Appli
43	126.6	8.7	1453	13	US-10-114-270-185	Sequence 185, App
44	113.6	7.8	1918	17	US-10-437-963-36340	Sequence 36340, A
45	108.8	7.5	3540	17	US-10-437-963-75672	Sequence 75672, A

#### ALIGNMENTS

#### RESULT 1

US-10-803-278-3  
; Sequence 3, Application US/10803278  
; Publication No. US20040175749A1  
; GENERAL INFORMATION:  
; APPLICANT: Walke, D. Wade  
; APPLICANT: Scoville, John  
; APPLICANT: Friddle, Carl Johan  
; TITLE OF INVENTION: Novel Human Kinases and Polynucleotides Encoding the Same  
; FILE REFERENCE: LEX-0348-USA  
; CURRENT APPLICATION NUMBER: US/10/803,278  
; CURRENT FILING DATE: 2004-03-18  
; PRIOR APPLICATION NUMBER: US/10/196,927  
; PRIOR FILING DATE: 2002-05-20  
; PRIOR APPLICATION NUMBER: US 60/293,248  
; PRIOR FILING DATE: 2001-05-24  
; NUMBER OF SEQ ID NOS: 5  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 3  
; LENGTH: 1449  
; TYPE: DNA  
; ORGANISM: homo sapiens  
US-10-803-278-3

Query Match	100.0%	Score	1449	DB	17	Length	1449
Best Local Similarity	100.0%	Pred. No.	0				
Matches	1449	Conservative	0	Mismatches	0	Indels	0
Gaps	0						
QY	1	ATGCTGAAATTC	CAAGAGGCAGCTTAAGTGTGTGAGTGATCAACAGCCATTTCACCTTAT	60			
Db	1	ATGCTGAAATTC	CAAGAGGCAGCTTAAGTGTGTGAGTGATCAACAGCCATTTCACCTTAT	60			
QY	61	CCAAAGACCTTGATTC	CAAGAGATACGTGCTTCAACAAAACTTGGCAGTGGAGTTT	120			
Db	61	CCAAAGACCTTGATTC	CAAGAGATACGTGCTTCAACAAAACTTGGCAGTGGAGTTT	120			
QY	121	GGAAGTGTCTATCTG	TTTCAGACAAAGAACCAACAGGAGGAGGAATTAAGGTACTT	180			

```

Db 121 GGAACCTGCTATCTGGTTTCAGACAAAGAGCCAAACGAGGAGGAATTTAAAGGTACTT 180
Qy 181 AAGGAAATATCTGTGTGAGAACTAAATCCAAATGAAACTGTACAGGCCAATTTGGAGCC 240
Db 181 AAGGAAATATCTGTGTGAGAACTAAATCCAAATGAAACTGTACAGGCCAATTTGGAGCC 240
Qy 241 CAACCTCTCCAGCTGAGCCACCCAGCCATTTGTCAAGTTCCATGCAAGTTTGTGGAG 300
Db 241 CAACCTCTCCAGCTGAGCCACCCAGCCATTTGTCAAGTTCCATGCAAGTTTGTGGAG 300
Qy 301 CAAGATAATTTCTGCATTTATCAGGAGTACTGTGAGGGCCGAGATCTGGACGATATAAAT 360
Db 301 CAAGATAATTTCTGCATTTATCAGGAGTACTGTGAGGGCCGAGATCTGGACGATATAAAT 360
Qy 361 CAGGATATATAAAGCTGGAAATATCTTTCCAGAAATCAAAATATAGAAATGATATC 420
Db 361 CAGGATATATAAAGCTGGAAATATCTTTCCAGAAATCAAAATATAGAAATGATATC 420
Qy 421 CAGCTGCTGGAGTGTGACTACATGCGATGAGGAGGATATCTTCATCGAGACTTAAAG 480
Db 421 CAGCTGCTGGAGTGTGACTACATGCGATGAGGAGGATATCTTCATCGAGACTTAAAG 480
Qy 481 TCAAGAAATCTATTTCTGAAATAATCTCTTTAAATTTGGAGTTTGGAGTTTCTGA 540
Db 481 TCAAGAAATCTATTTCTGAAATAATCTCTTTAAATTTGGAGTTTGGAGTTTCTGA 540
Qy 541 CTCTAATGGATCTGTGACCTGGCCACCACTTTAACTGGAACTCCCATTTATAGT 600
Db 541 CTCTAATGGATCTGTGACCTGGCCACCACTTTAACTGGAACTCCCATTTATAGT 600
Qy 601 CCTGAGCTCTGAAACACCAAGCTATGACACAAAGTCGACATCTGTCTACTGGCATGC 660
Db 601 CCTGAGCTCTGAAACACCAAGCTATGACACAAAGTCGACATCTGTCTACTGGCATGC 660
Qy 661 ATTTGTATGAGATGTGTCATGATATCATGATTCGCTGGCTCCATTTCTTATCCATT 720
Db 661 ATTTGTATGAGATGTGTCATGATATCATGATTCGCTGGCTCCATTTCTTATCCATT 720
Qy 721 GTTTTAAATTTCTGAAGTGACACACCTCTCTCCCTGAGAGATATCAAAAGACTA 780
Db 721 GTTTTAAATTTCTGAAGTGACACACCTCTCTCCCTGAGAGATATCAAAAGACTA 780
Qy 781 AATGCCATCATGAAAGCATGTGAAACAAAGATCTCTTCAATTAAGACCATCTGCTATCGAA 840
Db 781 AATGCCATCATGAAAGCATGTGAAACAAAGATCTCTTCAATTAAGACCATCTGCTATCGAA 840
Qy 841 ATTTTAAATTCCTTACCTTGATGAGCAGCTACAGACCTTAATGTGTAGATATTCAGAA 900
Db 841 ATTTTAAATTCCTTACCTTGATGAGCAGCTACAGACCTTAATGTGTAGATATTCAGAA 900
Qy 901 ATGACTCTGGAAGACAAAAATTTGATGTGAGAGGGCTGCTCATATATTAATGCC 960
Db 901 ATGACTCTGGAAGACAAAAATTTGATGTGAGAGGGCTGCTCATATATTAATGCC 960
Qy 961 ATGCATAAAGGATCCACCTGAGACTCTGAGGCACTGTGAGAGTACAGAAATGAGC 1020
Db 961 ATGCATAAAGGATCCACCTGAGACTCTGAGGCACTGTGAGAGTACAGAAATGAGC 1020
Qy 1021 CCAAGGAAAGGATGCGGCTGAGGAGCTCCAGGGGGCTGATGAGAAAGCCAGGAGCTG 1080
Db 1021 CCAAGGAAAGGATGCGGCTGAGGAGCTCCAGGGGGCTGATGAGAAAGCCAGGAGCTG 1080
Qy 1081 AAAAGATTTGGAAGAAAAATATGAGAAATATAGCAACGATTCGAAGATTTGAGATCT 1140
Db 1081 AAAAGATTTGGAAGAAAAATATGAGAAATATAGCAACGATTCGAAGATTTGAGATCT 1140
Qy 1141 CGGAATTTTTCAGAGCTGAGTGTGATGTACTTCCATGAAAAACACATTTAAAGGAATG 1200
Db 1141 CGGAATTTTTCAGAGCTGAGTGTGATGTACTTCCATGAAAAACACATTTAAAGGAATG 1200
Qy 1201 GAAGAAAGGAGAGCACTGAGGAGAGACTTTCTTGTTCACCCAGGACGAGGATGAA 1260

```

```

Db 1201 GAAGAAAAGGAGGAGCAACCTGAGGGAAGACTTTCTTTGTTTCACCCCAGGAGGATGAA 1260
Qy 1261 GAGAGTGGCAAGGAGGGAAGAGAAATCTGATGAACCAACTTTAGAGAACTTCCTGAG 1320
Db 1261 GAGAGTGGCAAGGAGGGAAGAGAAATCTGATGAACCAACTTTAGAGAACTTCCTGAG 1320
Qy 1321 TCTGAGCTATTTCTTCCATGAGACTCTCAAGAACTTGAATCAATTTAGAGGATGCGACA 1380
Db 1321 TCTGAGCTATTTCTTCCATGAGACTCTCAAGAACTTGAATCAATTTAGAGGATGCGACA 1380
Qy 1381 TCTGAGCTTGGATACCATGAGACTGTAAATCTAATTTCACTAGAGGATTTGGAAAAAT 1440
Db 1381 TCTGAGCTTGGATACCATGAGACTGTAAATCTAATTTCACTAGAGGATTTGGAAAAAT 1440
Qy 1441 GAAAAATAA 1449
Db 1441 GAAAAATAA 1449

RESULT 2
US-10-803-278-1
; Sequence 1, Application US/10803278
; Publication No. US2004017549A1
; GENERAL INFORMATION:
; APPLICANT: Walke, D. Wade
; APPLICANT: Scoville, John
; APPLICANT: Friddle, Carl Johan
; TITLE OF INVENTION: Novel Human Kinases and Polynucleotides Encoding the Same
; FILE REFERENCE: LEX-0348-USA
; CURRENT APPLICATION NUMBER: US/10/803,278
; CURRENT FILING DATE: 2004-03-18
; PRIOR APPLICATION NUMBER: US/10/196,927
; PRIOR FILING DATE: 2002-05-20
; PRIOR APPLICATION NUMBER: US 60/293,248
; PRIOR FILING DATE: 2001-05-24
; NUMBER OF SEQ ID NOS: 5
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 1938
; TYPE: DNA
; ORGANISM: homo sapiens
US-10-803-278-1

Query Match 96.5%; Score 1399; DB 17; Length 1938;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1399; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ATGCTGAAATTCAGAGGCGAGCTAAGTGTGTGAGTGGATCAACAGCCATTTCCATTAT 60
Db 1 ATGCTGAAATTCAGAGGCGAGCTAAGTGTGTGAGTGGATCAACAGCCATTTCCATTAT 60
Qy 61 CCAAGAACCTTGTATTCAGAGAGATACGCTCTTCAACAAAACTTGGCAGTGGAGTTT 120
Db 61 CCAAGAACCTTGTATTCAGAGAGATACGCTCTTCAACAAAACTTGGCAGTGGAGTTT 120
Qy 121 GGAACCTGTCTATCTGTTTCAGACAAAGAACCCAAACGAGGAGGAAATTAAGGTACTT 180
Db 121 GGAACCTGTCTATCTGTTTCAGACAAAGAACCCAAACGAGGAGGAAATTAAGGTACTT 180
Qy 181 AAGGAAATATCTGTGAGAACTAAATCCAAATGAAACTGTACAGGCCAATTTGGAGCC 240
Db 181 AAGGAAATATCTGTGAGAACTAAATCCAAATGAAACTGTACAGGCCAATTTGGAGCC 240
Qy 241 CAACCTCTCCAGCTGAGCCACCCAGCCATTTGTCAAGTTCCATGCAAGTTTGTGGAG 300
Db 241 CAACCTCTCCAGCTGAGCCACCCAGCCATTTGTCAAGTTCCATGCAAGTTTGTGGAG 300
Qy 301 CAAGATAATTTCTGCATTTATCAGGAGTACTGTGAGGGCCGAGATCTGGACGATATAAAT 360
Db 301 CAAGATAATTTCTGCATTTATCAGGAGTACTGTGAGGGCCGAGATCTGGACGATATAAAT 360
Qy 361 CAGGATATATAAAGCTGGAAATATCTTTCCAGAAATCAAAATATAGAAATGATGTTATC 420

```

```
Db 361 CAGGAATATAACAAGCTGGAAAAATCTTTCCAGAAAAATCAATAATAGATGGTTTATC 420
QY 421 CAGCTCTCTGCTGGAGTTGACTACATGATGAGGAGGAGGATCTTCATCGAGACTTAAG 480
Db 421 CAGCTCTCTGCTGGAGTTGACTACATGATGAGGAGGAGGATCTTCATCGAGACTTAAG 480
QY 481 TCAAGAAATGTATTTCTGAAAAATATATCTCTTAAATTTGGAGATTTTGGAGTTCTCGA 540
Db 481 TCAAGAAATGTATTTCTGAAAAATATATCTCTTAAATTTGGAGATTTTGGAGTTCTCGA 540
QY 541 CTTCTAATGGGATCTGTGACCTGGGCCACAACCTTTAACTGGGACTCCCATTTATGAT 600
Db 541 CTTCTAATGGGATCTGTGACCTGGGCCACAACCTTTAACTGGGACTCCCATTTATGAT 600
QY 601 CTTGAGGCTCTGAAACACCAAGCTATGACACAAAGTCGGACATCTGGTCACTGGCATGC 660
Db 601 CTTGAGGCTCTGAAACACCAAGCTATGACACAAAGTCGGACATCTGGTCACTGGCATGC 660
QY 661 ATTGTTGATGAGATGTGCTGCATGATCATGATTCGCTGGCTCCAAATTTCTTATCCAT 720
Db 661 ATTGTTGATGAGATGTGCTGCATGATCATGATTCGCTGGCTCCAAATTTCTTATCCAT 720
QY 721 GTTTTAAAAATTTTCAAGGTGACACACCTCTCTCCCTGAGAGATATCAAAAAGACTA 780
Db 721 GTTTTAAAAATTTTGAAGGTGACACACCTCTCTCCCTGAGAGATATCAAAAAGACTA 780
QY 781 ATGCCATCATGGAAGAGATGTTGAACAAGATCTTTCATTTAAGACCATCTGCTATCGAA 840
Db 781 AATGCCATCATGGAAGAGATGTTGAACAAGATCTTTCATTTAAGACCATCTGCTATCGAA 840
QY 841 ATTTTAAAAATCCCTTACCTTGTATGAGCAGCTACAGAACCTAATGTTAGATATTCAGAA 900
Db 841 ATTTTAAAAATCCCTTACCTTGTATGAGCAGCTACAGAACCTAATGTTAGATATTCAGAA 900
QY 901 ATGACTCTCGAAGACAAAAATTTGGATTTGTGAGAGGAGGCTGCTCATATAATTAATGCC 960
Db 901 ATGACTCTCGAAGACAAAAATTTGGATTTGTGAGAGGAGGCTGCTCATATAATTAATGCC 960
QY 961 ATGCAAAAAGGATCCACCTGCAGACTCTGAGGGCACTCTCAGAGTACAGAAAATGAGC 1020
Db 961 ATGCAAAAAGGATCCACCTGCAGACTCTGAGGGCACTCTCAGAGTACAGAAAATGAGC 1020
QY 1021 CCAAGAGAAAGGATGGCGGTGAGGAAGCTCCAGGGCGCTGATGAGAAAGCCAGGAAGCTG 1080
Db 1021 CCAAGAGAAAGGATGGCGGTGAGGAAGCTCCAGGGCGCTGATGAGAAAGCCAGGAAGCTG 1080
QY 1081 AAAAGATTTGGAAGAAAAATATGAAGAAAAATAGCAAAAGAAATGCAAAATGAGATCT 1140
Db 1081 AAAAGATTTGGAAGAAAAATATGAAGAAAAATAGCAAAAGAAATGCAAAATGAGATCT 1140
QY 1141 CGGAACCTTTGAGAGCTGAGTGTGATGATCTCATGAAAAACACATTTAAAGGAATG 1200
Db 1141 CGGAACCTTTGAGAGCTGAGTGTGATGATCTCATGAAAAACACATTTAAAGGAATG 1200
QY 1201 GAAGAAAGGAGGAGCAACTGAGGGAAGACTTTTCTGTTCACTCCAGGACGAGGATGAA 1260
Db 1201 GAAGAAAGGAGGAGCAACTGAGGGAAGACTTTTCTGTTCACTCCAGGACGAGGATGAA 1260
QY 1261 GAGAGTGGCAGGCGGAGGAGGAATCTGATGACCACTTTAGAGAACTGCTGCTGAG 1320
Db 1261 GAGAGTGGCAGGCGGAGGAGGAATCTGATGACCACTTTAGAGAACTGCTGCTGAG 1320
QY 1321 TCTCAGCTATTCTCTCCATGGACCTCCAGAACTTGAATCAATTTAGAGGATGCCACA 1380
Db 1321 TCTCAGCTATTCTCTCCATGGACCTCCAGAACTTGAATCAATTTAGAGGATGCCACA 1380
QY 1381 TCTGACCTTGGATACCATG 1399
Db 1381 TCTGACCTTGGATACCATG 1399
```

RESULT 3

US-10-333-314-25

```
; Sequence 25, Application US/10333314
; Publication No. US20030211093A1
; GENERAL INFORMATION:
; APPLICANT: INCYTE GENOMICS, INC.; YUE, Henry
; APPLICANT: KHAN, Farrah A.; GURURAJAN, Rajagopal
; APPLICANT: HAPALIA, April J.A.; CHAWLA, Narinder K.
; APPLICANT: ARVIZU, Chandra S.; RAMKUMAR, Jayalaxmi
; APPLICANT: GAUNDHI, Aneena R.; POLICKY, Jennifer L.
; APPLICANT: BAUGHN, Mariah R.; TRIBOULEY, Catherine M.
; APPLICANT: THORNTON, Michael B.; BANDMAN, Olga
; APPLICANT: NGUYEN, Danniell B.; LU, Yan
; APPLICANT: BURFORD, Neil; LAL, Preeti G.
; APPLICANT: DING, Li; YAO, Monique G.
; APPLICANT: ELLIOTT, Vicki S.; RECIPON, Shirley A.
; APPLICANT: KEARNEY, Liam; LU, Dyung Aina M.
; APPLICANT: GREENWALD, Sara R.; TANG, Y. Tom
; APPLICANT: XU, Yuming; WALSH, Roderick T.
; APPLICANT: GISTZEN, Kimberly J.; YANG, Junming
; APPLICANT: JACKSON, Jennifer L.
; TITLE OF INVENTION: HUMAN KINASES
; FILE REFERENCE: PI-0162 USN
; CURRENT APPLICATION NUMBER: US/10/333,314
; PRIOR FILING DATE: 2003-01-15
; PRIOR APPLICATION NUMBER: PCT/US01/23092
; PRIOR FILING DATE: 2001-07-20
; PRIOR APPLICATION NUMBER: US 60/220,038
; PRIOR FILING DATE: 2000-07-21
; PRIOR APPLICATION NUMBER: US 60/222,112
; PRIOR FILING DATE: 2000-07-28
; PRIOR APPLICATION NUMBER: US 60/222,831
; PRIOR FILING DATE: 2000-08-04
; PRIOR APPLICATION NUMBER: US 60/224,729
; NUMBER OF SEQ ID NOS: 40
; SOFTWARE: PERL Program
; SEQ ID NO 25
; LENGTH: 2552
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; OTHER INFORMATION: Incyte ID No: 7477585CB1
US-10-333-314-25
```

```
Query Match 96.5%; Score 1399; DB 13; Length 2552;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1399; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGCTGAAATTCGAAGAGGAGCTAAGTGTGTGAGTGAATCAACAGCCATTTCACCTTAT 60
Db 278 ATGCTGAAATTCGAAGAGGAGCTAAGTGTGTGAGTGAATCAACAGCCATTTCACCTTAT 337
QY 61 CCAAGACCTTGATGCAAGAGATAGTGTCTCAACAAAACTTGGCAGTGGAGTTT 120
Db 338 CCAAGACCTTGATGCAAGAGATAGTGTCTCAACAAAACTTGGCAGTGGAGTTT 397
QY 121 GGAAGTGTCTATCTGTTTCAGACAAGAAAGCCAAACGAGGAGAGGAATTAAGGTACTT 180
Db 398 GGAAGTGTCTATCTGTTTCAGACAAGAAAGCCAAACGAGGAGAGGAATTAAGGTACTT 457
QY 181 AAGGAATATCTGTGAGAACTAAATCCAAATGAATCTGACGGCCATTTGGAAGCC 240
Db 458 AAGGAATATCTGTGAGAACTAAATCCAAATGAATCTGACGGCCATTTGGAAGCC 517
QY 241 CAACCTCTCCAGCTGGACCCAGCCATTCGTAAGTTCCATGCAAGTTTGTGGAG 300
Db 518 CAATCTCTCCAGCTGGACCCAGCCATTCGTAAGTTCCATGCAAGTTTGTGGAG 577
QY 301 CAAGATAATTTCTGCATTATCAGGAGTACTGTGAGGCGGAGATCTGGACGATAAATT 360
Db 578 CAAGATAATTTCTGCATTATCAGGAGTACTGTGAGGCGGAGATCTGGACGATAAATT 637
QY 361 CAGGAATATAACAAGCTGGNAAAATCTTCCAGAAAATCAATAATAGATGGTTTATC 420
```



638 CAGAAATATAACACAGCTGGAAAAATCTTTCCAGAAAAATCAATAATAGAAATGTTTATC 697  
421 CAGCTGCTGCTGGAGTGTACATCATGATGAGGAGGATATCTTCATCGAGACTTAAAG 480  
698 CAGCTGCTGCTGGAGTGTACATCATGATGAGGAGGATATCTTCATCGAGACTTAAAG 757  
481 TCAAGAAATGATTTCTGAAAAATATCTCTTAAATTTGGAGATTTTGGAGTTTCTCGA 540  
758 TCAAGAAATGATTTCTGAAAAATATCTCTTAAATTTGGAGATTTTGGAGTTTCTCGA 817  
541 CTTCTAATGGATCTCTGTGACCTGGCCACAACTTTTAACCTGGAACTCCCATATATAGCT 600  
818 CTTCTAATGGATCTCTGTGACCTGGCCACAACTTTTAACCTGGAACTCCCATATATAGCT 877  
601 CTTGAGGCTCTGAAACACACAGGCTATGACAAAGTGGACATCTGGTCACTGGCATGC 660  
878 CTTGAGGCTCTGAAACACACAGGCTATGACAAAGTGGACATCTGGTCACTGGCATGC 937  
661 ATTTTGTATGAGATGCTGATGATCATGCAATTCGCTGGCTCCCAATTTCTTATCCATT 720  
938 ATTTTGTATGAGATGCTGATGATCATGCAATTCGCTGGCTCCCAATTTCTTATCCATT 997  
721 GTTTTAAAAATTTGTGAAGTGTACACACCTTTCTCTCCCTGAGAGATATCAAAAGAACTA 780  
998 GTTTTAAAAATTTGTGAAGTGTACACACCTTTCTCTCCCTGAGAGATATCAAAAGAACTA 1057  
781 AATGCCATCATGAAAGCATGTTGAACAGAAATCTTCTTAAAGACATCTGCTATCGAA 840  
1058 AATGCCATCATGAAAGCATGTTGAACAGAAATCTTCTTAAAGACATCTGCTATCGAA 1117  
841 ATTTTAAAAATTTGTGATGATGAGCTACAGAACTTAAATGTGATATTCAGAA 900  
1118 ATTTTAAAAATTTGTGATGATGAGCTACAGAACTTAAATGTGATATTCAGAA 1177  
901 ATGACTCTGAAACAAAAATTTGGATTTGCAAGAGGCTGCTCATATTAATGTC 960  
1178 ATGACTCTGAAACAAAAATTTGGATTTGCAAGAGGCTGCTCATATTAATGTC 1237  
961 ATGCAAAAAAGGATCCACCTGACAGCTCTGAGGGCACTGTGAGAGTACAGAAAAATGAG 1020  
1238 ATGCAAAAAAGGATCCACCTGACAGCTCTGAGGGCACTGTGAGAGTACAGAAAAATGAG 1297  
1021 CCAAGAAAGGATGCGGCTGAGGAAGCTCCAGCGGCTGATGAGAAACCGAGAGCTG 1080  
1298 CCAAGAAAGGATGCGGCTGAGGAAGCTCCAGCGGCTGATGAGAAACCGAGAGCTG 1357  
1081 AAAAAAGATTTCTGAAAGAAAAATATGAAGAAATAGCAAAACGAATGCAAGATTTGAGATCT 1140  
1358 AAAAAAGATTTCTGAAAGAAAAATATGAAGAAATAGCAAAACGAATGCAAGATTTGAGATCT 1417  
1141 CGGAATTTTCAGAGCTGAGTGTGATGATCTCCATGAAAAACACATTTAAAGGAATG 1200  
1418 CGGAATTTTCAGAGCTGAGTGTGATGATCTCCATGAAAAACACATTTAAAGGAATG 1477  
1201 GAAGAAAGAGGAGCACTGAGGAGAGACTTTCTTGTTCACCCAGGACGAGGATGAA 1260  
1478 GAAGAAAGAGGAGCACTGAGGAGAGACTTTCTTGTTCACCCAGGACGAGGATGAA 1537  
1261 GAGAGTGGCAAGCAGGGAAGGAGATCTGTGAAACCACTTTAGAGAACCTGCTGAG 1320  
1538 GAGAGTGGCAAGCAGGGAAGGAGATCTGTGAAACCACTTTAGAGAACCTGCTGAG 1597  
1321 TCTCAGCTATTCTTCCATGGAACCTCCAGAACTTGAATCAATTTAGAGGATGCCACA 1380  
1598 TCTCAGCTATTCTTCCATGGAACCTCCAGAACTTGAATCAATTTAGAGGATGCCACA 1657  
1381 TCTGACCTTGGATACCATG 1399  
1658 TCTGACCTTGGATACCATG 1676

US-10-425-114-22275  
; Sequence 22275, Application US/10425114  
; Publication No. US20040034888A1  
; GENERAL INFORMATION:  
; APPLICANT: Liu, Jingdong  
; APPLICANT: Zhou, Yihua  
; APPLICANT: Kovalic, David K.  
; APPLICANT: Screen, Steven E.  
; APPLICANT: Tabaska, Jack E.  
; APPLICANT: Cao, Yongwei  
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With  
; FILE OF INVENTION: Plants and Uses Thereof for Plant Improvement  
; FILE REFERENCE: 38-21(53313)B  
; CURRENT APPLICATION NUMBER: US/10/425,114  
; CURRENT FILING DATE: 2003-04-28  
; NUMBER OF SEQ ID NOS: 73128  
; SEQ ID NO 22275  
; LENGTH: 2669  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
; FEATURE:  
; OTHER INFORMATION: Clone ID: LIB3547-006-G1\_FLI  
US-10-425-114-22275

Query Match 96.5%; Score 1399; DB 13; Length 2669;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 1399; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGCTGAAATTCCTCAAGAGGAGCTAAGTGTGTGATCGATCAACAGCCATTTCCACTTAT 60  
DB 48 ATGCTGAAATTCCTCAAGAGGAGCTAAGTGTGTGATCGATCAACAGCCATTTCCACTTAT 107  
QY 61 CCAAGAGCTTTGATTTGCAAGAGATACGTCTTCAACAAAACTTGGCAGTGGAGTTT 120  
DB 108 CCAAGAGCTTTGATTTGCAAGAGATACGTCTTCAACAAAACTTGGCAGTGGAGTTT 167  
QY 121 GGAATCTCTATCTGTTTTCAGACAGAAAGCCAAACGAGGAGGAAATTAAGGTTACT 180  
DB 168 GGAATCTCTATCTGTTTTCAGACAGAAAGCCAAACGAGGAGGAAATTAAGGTTACT 227  
QY 181 AAGAAATATCTGTTGGAGAACTAAATCCAAATGAAACTGTACAGGCCAATTTGGAAGCC 240  
DB 228 AAGAAATATCTGTTGGAGAACTAAATCCAAATGAAACTGTACAGGCCAATTTGGAAGCC 287  
QY 241 CAATCTCTTCCAGCTGGACCCAGCCATTTGTCAGTTCCATGCAAGTTTGTGGAG 300  
DB 288 CAATCTCTTCCAGCTGGACCCAGCCATTTGTCAGTTCCATGCAAGTTTGTGGAG 347  
QY 301 CAAGATAATTTCTGCATTTACGCGAGTACTGTGAGGGCCGAGATCTGGACGATAAAAT 360  
DB 348 CAAGATAATTTCTGCATTTACGCGAGTACTGTGAGGGCCGAGATCTGGACGATAAAAT 407  
QY 361 CAGAAATATAACAGCTGGAAAAATCTTTCCAGAAAAATCAAAATATAGAAATGTTTATC 420  
DB 408 CAGAAATATAACAGCTGGAAAAATCTTTCCAGAAAAATCAAAATATAGAAATGTTTATC 467  
QY 421 CAGCTGCTGCTGGAGTTGATCATCATGAGAGGAGGATCTTTCATCGAGACTTAAAG 480  
DB 468 CAGCTGCTGCTGGAGTTGATCATCATGAGAGGAGGATCTTTCATCGAGACTTAAAG 527  
QY 481 TCAAGAAATGATTTCTGAAAAATTAATCTCTTAAAAATTTGGAGATTTTGGAGTTTCTCGA 540  
DB 528 TCAAGAAATGATTTCTGAAAAATTAATCTCTTAAAAATTTGGAGATTTTGGAGTTTCTCGA 587  
QY 541 CTTCTAATGGATCTCTGTGACCTGGCCACAACTTTTAACTGGAATCTCCCATATATAGCT 600  
DB 588 CTTCTAATGGATCTCTGTGACCTGGCCACAACTTTTAACTGGAATCTCCCATATATAGCT 647  
QY 601 CTTGAGGCTCTGAAACACACAGGCTATGACAAAGTGGACATCTGCTCACTGGCATGC 660  
DB 648 CTTGAGGCTCTGAAACACACAGGCTATGACAAAGTGGACATCTGCTCACTGGCATGC 707  
QY 661 ATTTTGTATGAGATGCTGCTGCAATCATGCAATTCGCTGGCTCCCAATTTCTTATCCATT 720



```

1168 ATGCAAAAAGGATCACCTGAGACTCTGAGGGCACTGTGAGAGTACAGAAATGACG 1227
1021 CCAAGAGAAAGGATCGGCTGAGGAAGCTCCAGGCGGCTGATGAGAAAGCCAGGAGCTG 1080
1228 CCAAGAGAAAGGATCGGCTGAGGAAGCTCCAGGCGGCTGATGAGAAAGCCAGGAGCTG 1287
1081 AAAAAAGTTGTGGAAGAAATATGAGAAATATGAGAAATATGAGAAATATGAGAAAT 1140
1288 AAAAAAGTTGTGGAAGAAATATGAGAAATATGAGAAATATGAGAAATATGAGAAAT 1347
1141 CGGAATTTTCAGCAGCTGAGTGTGATGATCTCCATGAAAGAAACACATTTAAAGAAATG 1200
1348 CGGAATTTTCAGCAGCTGAGTGTGATGATCTCCATGAAAGAAACACATTTAAAGAAATG 1407
1201 GAAGAAAGGAGGAGCAACCTGAGGAGAGGAGTCTTCTGTTTCAAGGAGGAGGATGAA 1260
1408 GAAGAAAGGAGGAGCAACCTGAGGAGAGGAGTCTTCTGTTTCAAGGAGGAGGATGAA 1467
1261 GAGAGTGGCAAGGAGGAGGAGGAGGAGTCTGATGAAAGGAGGAGTCTTCTGTTTCAAG 1320
1468 GAGAGTGGCAAGGAGGAGGAGGAGGAGTCTGATGAAAGGAGGAGTCTTCTGTTTCAAG 1527
1321 TCTCAGGCTATCTTCCATGAGGAGGAGGAGGAGTCTTCTGTTTCAAGGAGGAGGAGT 1380
1528 TCTCAGGCTATCTTCCATGAGGAGGAGGAGGAGTCTTCTGTTTCAAGGAGGAGGAGT 1587
1381 TCTGACCTTGGATACCATG 1399
1588 TCTGACCTTGGATACCATG 1606

```

RESULT 6

```

US-10-803-278-5
; Sequence 5: Application US/10803278
; Publication NO: US2004017549A1
; GENERAL INFORMATION:
; APPLICANT: Walke, D. Wade
; APPLICANT: Scoville, John
; APPLICANT: Friddle, Carl Johan
; TITLE OF INVENTION: Novel Human Kinases and Polynucleotides Encoding the Same
; FILE REFERENCE: LEX-0348-USA
; CURRENT APPLICATION NUMBER: US/10/803,278
; CURRENT FILING DATE: 2004-03-18
; PRIOR APPLICATION NUMBER: US/10/196,927
; PRIOR FILING DATE: 2002-05-20
; PRIOR APPLICATION NUMBER: US 60/293,248
; PRIOR FILING DATE: 2001-05-24
; NUMBER OF SEQ ID NOS: 5
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 5
; LENGTH: 2847
; TYPE: DNA
; ORGANISM: homo sapiens
US-10-803-278-5

```

```

Query Match 96.5%; Score 1399; DB 17; Length 2847;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1399; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGCTGAAATCCAGAGGAGGAGTAAAGTGTGAGTGGATCAACAGCCATTTCACCTTAT 60
DB 319 ATGCTGAAATCCAGAGGAGGAGTAAAGTGTGAGTGGATCAACAGCCATTTCACCTTAT 378
QY 61 CCAAGAGCTTGTATGAGAGAGATAGTGTCTTCAACAAAACCTTGGCAGTGGAGTTT 120
DB 379 CCAAGAGCTTGTATGAGAGAGATAGTGTCTTCAACAAAACCTTGGCAGTGGAGTTT 438
QY 121 GGAAGTGTCTATCTGTTTTCAGACAAGAAAGCCAAAGCAGGAGGAGGAAATTTAAAGGTACTT 180
DB 439 GGAAGTGTCTATCTGTTTTCAGACAAGAAAGCCAAAGCAGGAGGAGGAAATTTAAAGGTACTT 498
QY 181 AAGGAATATCTGTTGGAGAACTAAATCCAAATGTAAGAACTGTACAGGCCAATTTGGAAGCC 240

```

```

1499 AAGGAAATATCTGTGTGAGAACTAAATCCAAATGAAACTGTACAGGCCAAATTTGGAAGCC 558
241 CAATCTCTCCAGCTGGAGCCACCAGCCATTTGTCAAGTTCCATGCAAGTTTGTGGAG 300
559 CAATCTCTCCAGCTGGAGCCACCAGCCATTTGTCAAGTTCCATGCAAGTTTGTGGAG 618
301 CAAGATAATTTCTGCATTTATCAAGAGTACTGTGAGGGCCGAGATCTGAGACGATAAAAT 360
619 CAAGATAATTTCTGCATTTATCAAGAGTACTGTGAGGGCCGAGATCTGAGACGATAAAAT 678
361 CAGGAATATAAAGAGCTGGAAATCTTTCCAGAAATCAAAATATAGAAATGTTTATC 420
679 CAGGAATATAAAGAGCTGGAAATCTTTCCAGAAATCAAAATATAGAAATGTTTATC 738
421 CAGCTGCTGCTGGAGTGTGACTACATCATGAGAGGAGGATCTTTCATCGAGACTTAAAG 480
739 CAGCTGCTGCTGGAGTGTGACTACATCATGAGAGGAGGATCTTTCATCGAGACTTAAAG 798
481 TCAAGAGATGTTTCTGAAATATCTCTTAAATTTGAGATTTTGGAGTTTCTCGA 540
799 TCAAGAGATGTTTCTGAAATATCTCTTAAATTTGAGATTTTGGAGTTTCTCGA 858
541 CTTCTAAATGGGATCTGTGACCTGCGCCACAACTTTAACTGGAACCTCCCATTTATGAGT 600
859 CTTCTAAATGGGATCTGTGACCTGCGCCACAACTTTAACTGGAACCTCCCATTTATGAGT 918
601 CCGAGGCTCTGAAACACCAAGGCTATGACAAAGTCGGACATCTGCTCACTGGCATGC 660
919 CCGAGGCTCTGAAACACCAAGGCTATGACAAAGTCGGACATCTGCTCACTGGCATGC 978
661 ATTTTGTATGAGATGCTGTCATGAATCATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 720
979 ATTTTGTATGAGATGCTGTCATGAATCATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1038
721 GTTTTAAAAATTTGTGAAGTGTGACACACCTTCTCTCCCTGAGAGATATCCAAAGAACTA 780
1039 GTTTTAAAAATTTGTGAAGTGTGACACACCTTCTCTCCCTGAGAGATATCCAAAGAACTA 1098
781 AATGCCATCATGGAAGAGCTTTGAAGAAGATCTTCAATTAAGACCATCTGCTATCGAA 840
1099 AATGCCATCATGGAAGAGCTTTGAAGAAGATCTTCAATTAAGACCATCTGCTATCGAA 1158
841 ATTTTAAAAATTTGTGAAGTGTGACACACCTTCTCTCCCTGAGAGATATCCAAAGAACTA 900
1159 ATTTTAAAAATTTGTGAAGTGTGACACACCTTCTCTCCCTGAGAGATATCCAAAGAACTA 1218
901 ATGACTCTGGAAGAGCAAAAATTTGATGAGAGGAGGCTGCTCATATAATTAATGCC 960
1219 ATGACTCTGGAAGAGCAAAAATTTGATGAGAGGAGGCTGCTCATATAATTAATGCC 1278
961 ATGCAAAAAAGGATCTCACCTGAGAGTGTGAGGAGGAGTGTGAGAGGAGTGTGAGAGGAGT 1020
1279 ATGCAAAAAAGGATCTCACCTGAGAGTGTGAGGAGGAGTGTGAGAGGAGTGTGAGAGGAGT 1338
1021 CCAAGAGAAAGGATCTCGGCTGAGGAGCTCAGGCGGCTGATGAGAAAGCCAGGAGCTG 1080
1339 CCAAGAGAAAGGATCTCGGCTGAGGAGCTCAGGCGGCTGATGAGAAAGCCAGGAGCTG 1398
1081 AAAAGATTTGTGAAGAAATATGAGAAATATGAGAAATATGAGAAATATGAGAAATATGAGAA 1140
1399 AAAAGATTTGTGAAGAAATATGAGAAATATGAGAAATATGAGAAATATGAGAAATATGAGAA 1458
1141 CGGAATTTTCAGCAGCTGAGTGTGATGATCTCCATGAAAGAAACACATTTAAAGAAATG 1200
1459 CGGAATTTTCAGCAGCTGAGTGTGATGATCTCCATGAAAGAAACACATTTAAAGAAATG 1518
1201 GAAGAAAGGAGGAGCAACCTGAGGAGAGAGTCTTCTGTTTCAAGGAGGAGGAGGAGGAG 1260
1519 GAAGAAAGGAGGAGCAACCTGAGGAGAGAGTCTTCTGTTTCAAGGAGGAGGAGGAGGAG 1578
1261 GAGAGTGTGAGGAGGAGGAGGAGGAGGAGTCTGATGAACCAACTTTTAGAGAACTGCTGCTGAG 1320

```

Db 1579 GAGAGGTGGCAAGGCAAGGAGGAATCTGATGAACCAACTTTAGAGAACCTGCTGAG 1638  
Qy 1321 TCTCAGCCTATTCCTCCATGCGACCTCCAGAACTTGAATCAATTTAGAGATGCCACA 1380  
Db 1639 TCTCAGCCTATTCCTCCATGCGACCTCCAGAACTTGAATCAATTTAGAGATGCCACA 1698  
Qy 1381 TCTGACCTTGGATACCATG 1399  
Db 1699 TCTGACCTTGGATACCATG 1717

## RESULT 7

US-09-731-231A-1  
; Sequence 1, Application US/09731231A  
; Patent No. US20020082189A1  
; GENERAL INFORMATION:  
; APPLICANT: GUEGLER, Karl et al.  
; TITLE OF INVENTION: ISOLATED HUMAN KINASE PROTEINS, NUCLEIC  
; TITLE OF INVENTION: ACID MOLECULES ENCODING HUMAN KINASE PROTEINS, AND USES  
; TITLE OF INVENTION: THEREOF  
; FILE REFERENCE: CL001007  
; CURRENT APPLICATION NUMBER: US/09/731,231A  
; CURRENT FILING DATE: 2000-12-07  
; NUMBER OF SEQ ID NOS: 6  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 1  
; LENGTH: 2869  
; TYPE: DNA  
; ORGANISM: Human  
US-09-731-231A-1

Query Match 96.4%; Score 1397.4; DB 9; Length 2869;  
Best Local Similarity 99.9%; Pred. No. 0;  
Matches 1398; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
Qy 1 ATGCTGAAATTCACAGGAGCAGTAAAGTGTGAGTGGATCAACGCCATTTCCACTTAT 60  
Db 181 ATGCTGAAATTCACAGGAGCAGTAAAGTGTGAGTGGATCAACGCCATTTCCACTTAT 240  
Qy 61 CCAAGACCTTGATTCAGAGAGATACGCTTCAACAAAACCTTGGCAGTGGAAATTT 120  
Db 241 CCAAGACCTTGATTCAGAGAGATACGCTTCAACAAAACCTTGGCAGTGGAAATTT 300  
Qy 121 GGAAGCTGCTATCTGCTTTCAGACAGAAAGCCAAAGGAGGAGGAAATTAAGTACTT 180  
Db 301 GGAAGCTGCTATCTGCTTTCAGACAGAAAGCCAAAGGAGGAGGAAATTAAGTACTT 360  
Qy 181 AAGGAATATCTGTTGGAGAACTAAATCCAAATGAACTGTACAGGCCAAATTTGAAAGCC 240  
Db 361 AAGGAATATCTGTTGGAGAACTAAATCCAAATGAACTGTACAGGCCAAATTTGAAAGCC 420  
Qy 241 CAACTCTCTCCAGCTGACCAACCCAGCCATGTCAGTTCCATGCAAGTTTGTGGAG 300  
Db 421 CAACTCTCTCCAGCTGACCAACCCAGCCATGTCAGTTCCATGCAAGTTTGTGGAG 480  
Qy 301 CAAGATAATTTCTGCAATTCACGGAGTACTGTGAGGCGGAGATCTGACGATATAAAT 360  
Db 481 CAAGATAATTTCTGCAATTCACGGAGTACTGTGAGGCGGAGATCTGACGATATAAAT 540  
Qy 361 CAGGAATATAACAGCTGGAAAAATCTTTCCAGAAAAATCAAAATAGATGGTTTATC 420  
Db 541 CAGGAATATAACAGCTGGAAAAATCTTTCCAGAAAAATCAAAATAGATGGTTTATC 600  
Qy 421 CAGCTGCTCTGGAGTTCACTACATGATGAGAGGAGTACTTTCATCGAGCTTAAG 480  
Db 601 CAGCTGCTCTGGAGTTCACTACATGATGAGAGGAGTACTTTCATCGAGCTTAAG 660  
Qy 481 TCAAGAAATGATTTCTGAAAAATATCTCTTAAAAATGGAGATTTGGAGTTTCTCGA 540  
Db 661 TCAAGAAATGATTTCTGAAAAATATCTCTTAAAAATGGAGATTTGGAGTTTCTCGA 720  
Qy 541 CTTCTAATGGGATCTGTGAACTGGCCCAACTTAACTGGAACTCCCATATATAGT 600

Db 721 CTTCTAATGGGATCTGTGACCTGGCCACAACCTTAACTGGAACTCCCATATATAGT 780  
Qy 601 CTTGAGGCTCTGAAAACCAAGAGCTATGACAAAAGTCGACATCTGTCTACTGGCATGC 660  
Db 781 CTTGAGGCTCTGAAAACCAAGAGCTATGACAAAAGTCGACATCTGTCTACTGGCATGC 840  
Qy 661 ATTTTGTATGAGATGCTGTCATGATCAATGATGCTGCTGCTGCTGCTGCTGCTGCTGCT 720  
Db 841 ATTTTGTATGAGATGCTGTCATGATCAATGATGCTGCTGCTGCTGCTGCTGCTGCTGCT 900  
Qy 721 GTTTTAAAAAATGTTGAAGGTGACACACCTTCTCTCCCTGAGAGATATCCAAAAGAACTA 780  
Db 901 GTTTTAAAAAATGTTGAAGGTGACACACCTTCTCTCCCTGAGAGATATCCAAAAGAACTA 960  
Qy 781 AATGCCATCATGGAAGAGCATGTTGAACCAAGATCTTCAATTAAGACCATCTGCTATCGAA 840  
Db 961 AATGCCATCATGGAAGAGCATGTTGAACCAAGATCTTCAATTAAGACCATCTGCTATCGAA 1020  
Qy 841 ATTTTAAAAAATCCCTTACCTTGTGAGCAGCTACAGAACCTAAATGCTGATGATATTCAGAA 900  
Db 1021 ATTTTAAAAAATCCCTTACCTTGTGAGCAGCTACAGAACCTAAATGCTGATGATATTCAGAA 1080  
Qy 901 ATGACTCTGGAAGACAAAATTTGATGTCAGAGGAGGCTGCTCATATATTAATGCC 960  
Db 1081 ATGACTCTGGAAGACAAAATTTGATGTCAGAGGAGGCTGCTCATATATTAATGCC 1140  
Qy 961 ATGCAAAAAAGGATCCACCTGACAGCTCTGAGGCGACCTGTCAGAGTACAGAAAATGAGC 1020  
Db 1141 ATGCAAAAAAGGATCCACCTGACAGCTCTGAGGCGACCTGTCAGAGTACAGAAAATGAGC 1200  
Qy 1021 CCAAGAGAAAGGATCCGCTGAGGAAGCTCCAGGCGCTGATGAGAAAGCCAGGAAGCTG 1080  
Db 1201 CCAAGAGAAAGGATCCGCTGAGGAAGCTCCAGGCGCTGATGAGAAAGCCAGGAAGCTG 1260  
Qy 1081 AAAAGAGTTGTTGGAAGAAAATATGAGAAAATAGCAAAACGAAATGCAAGAAATTCAGATCT 1140  
Db 1261 AAAAGAGTTGTTGGAAGAAAATATGAGAAAATAGCAAAACGAAATGCAAGAAATTCAGATCT 1320  
Qy 1141 CGGAACCTTTCAGCAGCTGAGTGTGATGTCATCCATGAAAAAACAATTTAAAAAGGAATG 1200  
Db 1321 CGGAACCTTTCAGCAGCTGAGTGTGATGTCATCCATGAAAAAACAATTTAAAAAGGAATG 1380  
Qy 1201 GAAAGAAAGGAGGAGCACTGAGGGAAGACTTTCTGTTCCACCCAGGAGGAGATGAA 1260  
Db 1381 GAAAGAAAGGAGGAGCACTGAGGGAAGACTTTCTGTTCCACCCAGGAGGAGATGAA 1440  
Qy 1261 GAGAGGTGGCAAGCAGGAGAGGAAATCTGATGAACCACTTTAGAGAACCTGCTGAG 1320  
Db 1441 GAGAGGTGGCAAGCAGGAGAGGAAATCTGATGAACCACTTTAGAGAACCTGCTGAG 1500  
Qy 1321 TCTCAGCCTATTCCTTCAATGAGCTCCAGAACTTGAATCAATTTGAGAGGATGCCACA 1380  
Db 1501 TCTCAGCCTATTCCTTCAATGAGCTCCAGAACTTGAATCAATTTGAGAGGATGCCACA 1560  
Qy 1381 TCTCAGCCTTGGATACCATG 1399  
Db 1561 TCTCAGCCTTGGATACCATG 1579

## RESULT 8

US-10-751-985-1  
; Sequence 1, Application US/10751985  
; Publication No. US20040126861A1  
; GENERAL INFORMATION:  
; APPLICANT: GUEGLER, Karl et al.  
; TITLE OF INVENTION: ISOLATED HUMAN KINASE PROTEINS, NUCLEIC  
; TITLE OF INVENTION: ACID MOLECULES ENCODING HUMAN KINASE PROTEINS, AND USES  
; TITLE OF INVENTION: THEREOF  
; FILE REFERENCE: CL001007CON  
; CURRENT APPLICATION NUMBER: US/10/751,985  
; CURRENT FILING DATE: 2004-01-07  
; NUMBER OF SEQ ID NOS: 6  
; SOFTWARE: FastSeq for Windows Version 4.0

```

; SEQ ID NO 1
; LENGTH: 2869
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-751-985-1

Query Match      96.4%; Score 1397.4; DB 17; Length 2869;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 1398; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 ATGCTGAAATTCACAGGAGCAGTAAAGTGTGTGAGTGTGATCAACAGCCATTTCCACTTAT 60
DB 181 ATGCTGAAATTCACAGGAGCAGTAAAGTGTGTGAGTGTGATCAACAGCCATTTCCACTTAT 240
QY 61 CCAAGACCTTGATTCGAAGAGATACGTGCTTCAACAAAACCTTGGCAGTGTGAAATTT 120
DB 241 CCAAGACCTTGATTCGAAGAGATACGTGCTTCAACAAAACCTTGGCAGTGTGAAATTT 300
QY 121 GGAACCTGTCTATCTGTTTCAGACAAGAAAGCCAAACGAGGAGAGGAATTAAGGTACTT 180
DB 301 GGAACCTGTCTATCTGTTTCAGACAAGAAAGCCAAACGAGGAGAGGAATTAAGGTACTT 360
QY 181 AAGGAATATCTGTTGGAGAACTAAATCCAAATGAACTGTACAGGCCCAATTTGGAAGCC 240
DB 361 AAGGAATATCTGTTGGAGAACTAAATCCAAATGAACTGTACAGGCCCAATTTGGAAGCC 420
QY 241 CAATCTCTCCAAAGCTGGACCAACCCAGCCATTTGTCAAGTTTCCATGCAAGTTTGTGGAG 300
DB 421 CAATCTCTCTCCAAAGCTGGACCAACCCAGCCATTTGTCAAGTTTCCATGCAAGTTTGTGGAG 480
QY 301 CAAGATAATTTCTGCATATTCACAGGAGTACTGTGAGGCGGAGATCTGGACGATAAAAT 360
DB 481 CAAGATAATTTCTGCATATTCACAGGAGTACTGTGAGGCGGAGATCTGGACGATAAAAT 540
QY 361 CAGGAATATAAACAAGCTGAAAAATCTTCCAGAAAAATCAAAATATAAGTGTGTTATC 420
DB 541 CAGGAATATAAACAAGCTGAAAAATCTTCCAGAAAAATCAAAATATAAGTGTGTTATC 600
QY 421 CAGCTGCTGTGGAGTTGACTACATGATGAGAGAGGATATCTTCAATGAGACTTAAAG 480
DB 601 CAGCTGCTGTGGAGTTGACTACATGATGAGAGAGGATATCTTCAATGAGACTTAAAG 660
QY 481 TCAAGAGATGTTTCTGAAAAATTAATCTCTTAAAAATTCGAGATTTGGAGTTTCTCGA 540
DB 661 TCAAGAGATGTTTCTGAAAAATTAATCTCTTAAAAATTCGAGATTTGGAGTTTCTCGA 720
QY 541 CTTCTAATGGGATCCTGTGACCTGGCCACAACTTTAACTGGAACTCCCAATATATAGT 600
DB 721 CTTCTAATGGGATCCTGTGACCTGGCCACAACTTTAACTGGAACTCCCAATATATAGT 780
QY 601 CTTGAGGCTGTGAACACCAAGGATGACACAAAGTCGACATCTGTCACCTGGGATGC 660
DB 781 CTTGAGGCTGTGAACACCAAGGATGACACAAAGTCGACATCTGTCACCTGGGATGC 840
QY 661 ATTTTGTATGAGATGCTGATCAATCAATGATGCTGCTGCTCCCAATTTCTTATCCAT 720
DB 841 ATTTTGTATGAGATGCTGATCAATCAATGATGCTGCTGCTCCCAATTTCTTATCCAT 900
QY 721 GTTTTAAAAATTTGTAAGGTGACACACCTTCTCTCCCTGAGAGATATCAAAAGACTA 780
DB 901 GTTTTAAAAATTTGTAAGGTGACACACCTTCTCTCCCTGAGAGATATCAAAAGACTA 960
QY 781 AATGCCATCATGGAAGGATGTTGACAGAGATCTTCAATTAAGCACTCTGCTATCGAA 840
DB 961 AATGCCATCATGGAAGGATGTTGACAGAGATCTTCAATTAAGCACTCTGCTATCGAA 1020
QY 841 ATTTTAAAAATTCCTTACCTTACCTGATGAGAGGATGACAGAACTTAATGTGTAGATTCGAA 900
DB 1021 ATTTTAAAAATTCCTTACCTTACCTGATGAGAGGATGACAGAACTTAATGTGTAGATTCGAA 1080
QY 901 ATGACTCTGGAGACAAAAATTTGGATGTCAGAGGAGGCTGCTCATATAATTAATGCC 960
DB 1081 ATGACTCTGGAGACAAAAATTTGGATGTCAGAGGAGGCTGCTCATATAATTAATGCC 1140

```

```

RESULT 9
US-10-302-172-50
; Sequence 50, Application US/10302172
; Publication No. US20040053250A1
; GENERAL INFORMATION:
; APPLICANT: Tang, Y. Tom
; APPLICANT: Xue, Aidong J.
; APPLICANT: Drmanac, Radje T.
; TITLE OF INVENTION: No. US20040053250A1 Arginine-rich Protein-like Nucleic Acids and
; FILE REFERENCE: 803 1CNC
; CURRENT APPLICATION NUMBER: US/10/302,172
; CURRENT FILING DATE: 2002-11-21
; PRIOR APPLICATION NUMBER: US 10/225,251
; PRIOR FILING DATE: 2002-08-20
; PRIOR APPLICATION NUMBER: PCT US02/05095
; PRIOR FILING DATE: 2002-03-05
; PRIOR APPLICATION NUMBER: US 09/799,451
; PRIOR FILING DATE: 2001-03-05
; NUMBER OF SEQ ID NOS: 950
; SOFTWARE: pc_FL_genes Version 2.0
; SEQ ID NO 50
; LENGTH: 2765
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (597)..(2087)
US-10-302-172-50

```

```

Query Match      95.6%; Score 1385.4; DB 13; Length 2765;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 1397; Conservative 0; Mismatches 1; Indels 1; Gaps 1;

QY 1 ATGCTGAAATTCAGAGGAGCAGTAAAGTGTGTGAGTGTGATCAACAGCCATTTCCACTTAT 60
DB 154 ATGCTGAAATTCAGAGGAGCAGTAAAGTGTGTGAGTGTGATCAACAGCCATTTCCACTTAT 213
QY 61 CCAAGACCTTGTATTGCAAGAGATACGTTGCTTCAACAAAACTTGGCAGTGTGAAGTTTT 120

```

Db 214 CCAAGACCTTGATTCGAAGAGATACGCTCTTCAACAAATACTTGGCAGTGGAGTTT 273  
Qy 121 GGAACGTCTATCTGTTTCAGACAAGAACCCAGAGAGAGAAATTAAGGTACTT 180  
Db 274 GGAACGTCTATCTGTTTCAGACAAGAACCCAGAGAGAGAAATTAAGGTACTT 333  
Qy 181 AAGGAATATCTCTGAGAGAACTTAATCCAAATGAAACTGTACAGGCCAATTTGGAGCC 240  
Db 334 AAGGAATATCTCTGAGAGAACTTAATCCAAATGAAACTGTACAGGCCAATTTGGAGCC 393  
Qy 241 CAACCTCTCTCAAGCTGGACCCAGCCAGCCATTGTCAAGTTCACATGCAAG-TTGTGGAG 300  
Db 394 CAACCTCTCTCAAGCTGGACCCAGCCAGCCATTGTCAAGTTCACATGCAAG-TTGTGGAG 452  
Qy 301 CAGATAATTTCTGCATTATACGAGAGTACTGTGAGGCCGAGATCTGACGATTAATTT 360  
Db 453 CAGATAATTTCTGCATTATACGAGAGTACTGTGAGGCCGAGATCTGACGATTAATTT 512  
Qy 361 CAGGAATATAAACAAGCTGGAAAAATCTTTCCAGAAAAATCAATAATAGATGGTTATC 420  
Db 513 CAGGAATATAAACAAGCTGGAAAAATCTTTCCAGAAAAATCAATAATAGATGGTTATC 572  
Qy 421 CAGCTGCTGCTGGAGTGAATCATGCTATGAGAGAGGATCTTCATCGAGACTTAAAG 480  
Db 573 CAGCTGCTGCTGGAGTGAATCATGCTATGAGAGAGGATCTTCATCGAGACTTAAAG 632  
Qy 481 TCAAGAATATGTTTCTGAAAAATATCTCTTAAATTTGGAGATTTTGGAGTTTCTCGA 540  
Db 633 TCAAGAATATGTTTCTGAAAAATATCTCTTAAATTTGGAGATTTTGGAGTTTCTCGA 592  
Qy 541 CTTCTAATGGATTCCTGTGACCTGGCCACAACTTTAACTGGAACTCCCATTTATGAT 600  
Db 693 CTTCTAATGGATTCCTGTGACCTGGCCACAACTTTAACTGGAACTCCCATTTATGAT 752  
Qy 601 CTTGAGCTCTGAACCAAGCTATGACAAAGTTCGACATCTGTCACCTGGCATCC 560  
Db 753 CTTGAGCTCTGAACCAAGCTATGACAAAGTTCGACATCTGTCACCTGGCATCC 812  
Qy 661 ATTTCGTATGAGATGTGCTGATGATCATGATTCGCTGGCTCCCAATTTCTTATCCAT 720  
Db 813 ATTTCGTATGAGATGTGCTGATGATCATGATTCGCTGGCTCCCAATTTCTTATCCAT 872  
Qy 721 GTTTTAAAAATTTGAGAGTGACACACCTTCTCTCCCTGAGAGATATCCAAAAAGACTA 780  
Db 873 GTTTTAAAAATTTGAGAGTGACACACCTTCTCTCCCTGAGAGATATCCAAAAAGACTA 932  
Qy 781 AATGCCATCATGGAAGAGCTTTGAAACAAGATCTTCAATTAAGACCATCTGCTATCGAA 840  
Db 933 AATGCCATCATGGAAGAGCTTTGAAACAAGATCTTCAATTAAGACCATCTGCTATCGAA 992  
Qy 841 ATTTTAAAAATTCCTTACCTTGATGAGAGCTACAGAACTTAATGTATGATTTACAGAA 900  
Db 993 ATTTTAAAAATTCCTTACCTTGATGAGAGCTACAGAACTTAATGTATGATTTACAGAA 1052  
Qy 901 ATGACTCTGGAGACAAAAATTTGGATTTCTCAGAGAGGCTGCTCATATATTAATGCC 960  
Db 1053 ATGACTCTGGAGACAAAAATTTGGATTTCTCAGAGAGGCTGCTCATATATTAATGCC 1112  
Qy 961 ATGCAAAAAAGGATCCACCTGCAGACTCTGAGGGCACTCTCAGAGTACAGAAAAATGAGC 1020  
Db 1113 ATGCAAAAAAGGATCCACCTGCAGACTCTGAGGGCACTCTCAGAGTACAGAAAAATGAGC 1172  
Qy 1021 CCAAGAGAAAGGATGGGCTGAGGAAGCTCCAGGGCTGATGAGAAAGCCAGGAGCTG 1080  
Db 1173 CCAAGAGAAAGGATGGGCTGAGGAAGCTCCAGGGCTGATGAGAAAGCCAGGAGCTG 1232  
Qy 1081 AAAAGATTTGTGAAGAAAAATATGAAGAAAAATAGCAAAAGCAATGCAAGATTTGAGATCT 1140  
Db 1233 AAAAGATTTGTGAAGAAAAATATGAAGAAAAATAGCAAAAGCAATGCAAGATTTGAGATCT 1292  
Qy 1141 CGGAACTTTTCAGAGCTGAGTGTGATGTACTCCATGAAAAAACACATTTAAAAAGGATG 1200

Db 1293 CGGAATTTTCAGCAGCTGAGTGTGTGATGTACTCTCATGAAAAAACACATTTAAAGGAATG 1352  
Qy 1201 GAAGAAAAAGGAGGAGCAACCTGAGGGAAGACTTTCTTGTTCACCCAGAGAGGATGAA 1260  
Db 1353 GAAGAAAAAGGAGGAGCAACCTGAGGGAAGACTTTCTTGTTCACCCAGAGAGGATGAA 1412  
Qy 1261 GAGAGGTGCGAAGGAGGAGGAGGAAATCTGATGAACCAACTTTTAGAGAACTGCTCTGAG 1320  
Db 1413 GAGAGGTGCGAAGGAGGAGGAGGAAATCTGATGAACCAACTTTTAGAGAACTGCTCTGAG 1472  
Qy 1321 TCTCAGCCTATTTCCTTCCATGAGCCTCCACGAACCTTGAATCAATTTGTAGAGGATGCCACA 1380  
Db 1473 TCTCAGCCTATTTCCTTCCATGAGCCTCCACGAACCTTGAATCAATTTGTAGAGGATGCCACA 1532  
Qy 1381 TCTGACCTTTGGATACCATG 1399  
Db 1533 TCTGACCTTTGGATACCATG 1551

RESULT 10  
US-10-425-114-16420  
; Sequence 16420, Application US/10425114  
; Publication No. US20040034888A1  
; GENERAL INFORMATION:  
; APPLICANT: Liu, Jingdong  
; APPLICANT: Zhou, Yihua  
; APPLICANT: Kovalic, David K.  
; APPLICANT: Screen, Steven E.  
; APPLICANT: Tabaska, Jack E.  
; APPLICANT: Cao, Yongwei  
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With  
; FILE OF INVENTION: Plants and Uses Thereof for Plant Improvement  
; FILE REFERENCE: 38-21(53313)B  
; CURRENT APPLICATION NUMBER: US/10/425,114  
; CURRENT FILING DATE: 2003-04-28  
; NUMBER OF SEQ ID NOS: 73128  
; SEQ ID NO 16420  
; LENGTH: 2484  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
; FEATURE:  
; OTHER INFORMATION: Clone ID: LIB3063-070-A3\_FLI  
US-10-425-114-16420

Query Match 52.4%; Score 759; DB 13; Length 2484;  
Best Local Similarity 77.5%; Pred. No. 1e-205;  
Matches 1084; Conservative 0; Mismatches 0; Indels 315; Gaps 1;

Qy 1 ATGCTGAAATTCGAAGGAGGAGCTAAGTGTGTGAGTGGATCAACAGCCATTTCCACTTAT 60  
Db 267 ATGCTGAAATTCGAAGGAGGAGCTAAGTGTGTGAGTGGATCAACAGCCATTTCCACTTAT 326  
Qy 61 CCAAGACCTTTGATTGCAAGAAGATACGTGTTCAACAAAACTTGGCAGTGAAGTTT 120  
Db 327 CCAAGACCTTTGATTGCAAGAAGATACGTGTTCAACAAAACTTGGCAGTGAAGTTT 386  
Qy 121 GGAATCTCTATCTGTTTCAGACAAGAACCCAAATGAAATCTGACAGGCCAATTTGGAAGCC 180  
Db 387 GGAATCTCTATCTGTTTCAGACAAGAACCCAAATGAAATCTGACAGGCCAATTTGGAAGCC 446  
Qy 181 AAGGAATATCTGTTGGAGAACTAAATCCAAATGAAATCTGACAGGCCAATTTGGAAGCC 240  
Db 447 AAGGAATATCTGTTGGAGAACTAAATCCAAATGAAATCTGACAGGCCAATTTGGAAGCC 506  
Qy 241 CAATCTCTCTCAAGCTGGACCCAGCCAGCCATTGTCAAGTTCCATGCAAGTTTGTGAG 300  
Db 507 CAATCTCTCTCAAGCTGGACCCAGCCAGCCATTGTCAAGTTCCATGCAAGTTTGTGAG 566  
Qy 301 CAGATAATTTCTGCATTATCAGGAGTACTGTGAGGCCGAGATCTGAGAGGATTAATTT 360  
Db 567 CAGATAATTTCTGCATTATCAGGAGTACTGTGAGGCCGAGATCTGAGAGGATTAATTT 626  
Qy 361 CAGGAATATAAACAAGCTGGAAAAATCTTTCCAGAAAAATCAATAATAGATGGTTTATC 420







Db 1111 ATTTTATACAGATGTGTGTCGATCATCGTTTCTTGGCTCCAAATTTCTTGTCTGTG 1170  
Qy 721 GTTTTAAATTTGTTGAAGGTGACACACCTTCTCTCCCTGAGAGATATCCAAAGAACTA 780  
Db 1171 GTTTTGAATATCGTTGAAGGCAACACGCTTCACTCCCTGACAGATACCCAGAACTA 1230  
Qy 781 AATGCCATCATGGAAGCATGTTGAACAGAAATCTTTCATTAAGACCATCTGCTATCGAA 840  
Db 1231 AACATCATCATGTCACGCATGTTGAACAGAGCTCCCTCCCTGAGACCGCTCGCTGCAGAG 1290  
Qy 841 ATTTTAAATCCCTTACCTTCAATGAGCAGCTGACAGAACCTAATGTGTAGATATTCAGAA 900  
Db 1291 ATTTTAAAGCCCTTATGTGAAGAGTGCCTTCAGCAGCTGATGATATAAATCATCCAGAG 1350  
Qy 901 ATGACTCTCGAAGACAA---AAATTGGATTGTCAGAGAGGCTGCTCATATAATTAAT 957  
Db 1351 GCGACACTGGAAGACNAGAGGAACTCAGCTCTGACAGAGAGGCTGCCCATGCAGTTAAC 1410  
Qy 958 GCCATGCAAA 967  
Db 1411 GCCATAAAGA 1420

RESULT 12  
US-10-152-661-511  
; Sequence 511, Application US/10152661  
; Publication No. US20030022835A1  
; GENERAL INFORMATION:  
; APPLICANT: Watson, James D.  
; APPLICANT: Strachan, Lorna  
; APPLICANT: Sleeman, Matthew  
; APPLICANT: Onrust, Rene  
; APPLICANT: Murison, James G.  
; APPLICANT: Kumble, Krishanand D.  
; TITLE OF INVENTION: Compositions Isolated From Skin Cells  
; FILE REFERENCE: 11000.1011c5  
; CURRENT APPLICATION NUMBER: US/10/152,661  
; PRIOR FILING DATE: 2002-05-20  
; PRIOR APPLICATION NUMBER: 09/866,050  
; PRIOR FILING DATE: 2001-05-24  
; PRIOR APPLICATION NUMBER: 60/221,232  
; PRIOR FILING DATE: 2000-07-25  
; PRIOR APPLICATION NUMBER: 60/206,650  
; PRIOR FILING DATE: 2000-05-24  
; PRIOR APPLICATION NUMBER: 09/312,283  
; PRIOR FILING DATE: 1999-05-14  
; PRIOR APPLICATION NUMBER: PCT/NZ99/00051  
; PRIOR FILING DATE: 1999-04-29  
; PRIOR APPLICATION NUMBER: 09/188,930  
; PRIOR FILING DATE: 1998-11-09  
; PRIOR APPLICATION NUMBER: 09/069,726  
; PRIOR FILING DATE: 1998-04-29  
; NUMBER OF SEQ ID NOS: 725  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 511  
; LENGTH: 2538  
; TYPE: DNA  
; ORGANISM: Rat  
US-10-152-661-511

Query Match 46.2%; Score 669.6; DB 15; Length 2538;  
Best Local Similarity 82.2%; Pred. No. 3.8e-180;  
Matches 797; Conservative 0; Mismatches 164; Indels 9; Gaps 2;  
Qy 1 ATGCTGAAATCCAAAGGAGCAGTAAGTGTGTGAGTGATCAACAGCCATTTCCACTTAT 60  
Db 457 ATGCTGAAATCCAAAGAGAGCTGTAAGT-----GTGGACCCACAGCCATTTCTGTGTAC 510  
Qy 61 CCAAGACCTTGATTCAGAGAGATACGCTTCAACAAATTTGGCAGTGGAAAGTTT 120  
Db 511 CCACCCACTGTGATTCAGAGAGATACGCTTCCCAACAGAGCTTGGCAGCGGAGTTT 570

Qy 121 GGAACCTCTATCTTGGTTTCAGACAAGAAAGCCAAACGAGGAGAGGAAATTTAAAGGTACTT 180  
Db 571 GGAACCTCTATCTTGGTTTCAGACAAGAAAGCCAAACGATGAGAGGAACTTAAAGTACTG 630  
Qy 181 AAGAAATATCTGTGTGAGAACTTAATCCAAATGAAATCTGTACAGGCCAATTTGGAAGCC 240  
Db 631 AAGAAATATCTGTGTGAGAAATTAATCCAAATGAAATCTGTACAGGCCAATTTGGAAGCC 690  
Qy 241 CAACCTCTCTCCAAAGCTGGACCCAGCCATTTCAAGTTCCATGCAAGTTTGTGTGGAG 300  
Db 691 CAGCTCTCTCCAAAGCTGGACCCAGCCATTTCAAGTTCCATGCAAGTTTGTGTGGAG 750  
Qy 301 CAAGATTAATTTTGTGCAATTAACGGAGTACTGTGAGGCGCGAGATCTGGAAGCAATAAAT 360  
Db 751 CAGGTAATTTTGTGCAATTAACGGAGTACTGTGAGGCGCGAGATCTGGAAGCAATAAAT 810  
Qy 361 CAGCAATATAAACAAGCTGGAAATCTTTCCAGAAATCAATTAATAGAAATGTTTATC 420  
Db 811 CAGCAATATAAACAAGCTGGAAATCTTTCCAGAAATCAATTAATAGAAATGTTTATC 870  
Qy 421 CAGCTGCTGTGGAGTTGACTACATCATGAGGAGGAGATCTTTCATGAGACTTTAAAG 480  
Db 871 CAGTTGCTGTGGAGTTGACTACATCATGAGGAGGAGATCTTTCATGAGACTTTGAAA 930  
Qy 481 TCAAGAATGTAATTTCTGAAAAATTAATCTCTTAAATTTGAGATTTTGGAGTTTCTCGA 540  
Db 931 TCAAGAATGTAATTTCTGAAAAATTAATCTCTTAAATTTGAGATTTTGGAGTTTCTCGG 990  
Qy 541 CTCTAATGGGATCCTGTGACCTGGCCACAACTTTAACTGGAACCTCCCATTTATATGAGT 600  
Db 991 CTGCTGATGGTTCAATGTGAGCTGGCTACAACTTAACCGGGACTCCCATTTATATGAGT 1050  
Qy 601 CTTGAGCTCTGAAACACCAAGGCTATGACAAAGTCGAGACATCTGTGTCATGGAATGC 660  
Db 1051 CCAGAGCCCTGAAAGCACCAGGCTGTGATGCCAAAGTCTGACATCTGTCTACTGGCATGC 1110  
Qy 661 ATTTTGTATGAGATGCTGTCATGAATCATGCAATCGCTGCTCCCAATTTCTTATCCATT 720  
Db 1111 ATTTTACGAGATGTTGATGAGATCAATGCGTTTCTGGCTCCAATTTCTTCTGTG 1170  
Qy 721 GTTTTAAAAATTTGAAAGGTGACACACCTTCTCTCCCTGAGAGATATCCAAAGAACTA 780  
Db 1171 GTTTTGAATATCGTTGAAAGGCAACACGCTTCACTCCCTGACAGATACCCACGAGAACTA 1230  
Qy 781 AATGCCATCATGGAAGAGATGTTGAACAGAAATCTTTCATTAAGACCATCTGCTATCGAA 840  
Db 1231 AACATCATCATGTCACGCATGTTGAACAGAGTCCCTCCCTGAGACCGCTCGCTGCAGAG 1290  
Qy 841 ATTTTAAAAATCCCTTACCTTGTATGAGCAGCTACAGAAACCTTAATGTGTAGATATTCAGAA 900  
Db 1291 ATTTTAAAGCCCTTATGTGGAAGAGTGCCTTCAGCAGCTGATGATATAAATCAATCCAGAG 1350  
Qy 901 ATGACTCTGGAAGACAA---AAATTGGATTGTCAGAGAGGCTGCTCATATAATTAAT 957  
Db 1351 GCGACACTGGAAGACAAAGAGAACTCAGCTGTGCAAGAGGAGGCTGCCCATGCAGTTAAC 1410  
Qy 958 GCCATGCAAA 967  
Db 1411 GCCATAAAGA 1420

RESULT 13  
US-09-764-875-60  
; Sequence 60, Application US/09764875  
; Publication No. US20040018969A1  
; GENERAL INFORMATION:  
; APPLICANT: Rosen et al.  
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies  
; FILE REFERENCE: P0202  
; CURRENT APPLICATION NUMBER: US/09/764,875  
; CURRENT FILING DATE: 2001-01-17  
; Prior application data removed - consult PALM or file wrapper  
; NUMBER OF SEQ ID NOS: 1249

```
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 60
; LENGTH: 451
; TYPE: DNA
; ORGANISM: Homo sapiens
; NAME/KEY: SITE
; LOCATION: (307)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: SITE
; LOCATION: (363)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: SITE
; LOCATION: (368)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: SITE
; LOCATION: (392)
; OTHER INFORMATION: n equals a,t,g, or c
; OTHER INFORMATION: n equals a,t,g, or c
; US-09-764-875-60

Query Match      28.7%; Score 415.8; DB 11; Length 451;
Best Local Similarity 98.2%; Pred. No. 5.1e-108;
Matches 439; Conservative 0; Mismatches 6; Indels 2; Gaps 2;

Qy 738 AGTGACACACCTTCTCTCCCTGAGAGATATCCAAAGAACTAAATGCCATCATGGAAG 797
Db 2 AGTGACACACCTTCTCTCCCTGAGAGATATCCAAAGAACTAAATGCCATCATGGAAG 61

Qy 798 CATGTTGAACAAGAACCTTTCATTAAGACCATCTGCTATCGAAATTTTAAATAATCCCTTA 857
Db 62 CATGTTGAACAAGAACCTTTCATTAAGACCATCTGCTATCGAAATTTTAAATAATCCCTTA 121

Qy 858 CTTGTATGACGCTACAGAACCTTAATGTGTAGATATTCAGAAATGACTCTCGAAGACAA 917
Db 122 CTTGTATGACGCTACAGAACCTTAATGTGTAGATATTCAGAAATGACTCTCGAAGACAA 181

Qy 918 AAATTTGGATTGTGCAAGAGGCTCTCATATTAATTCATTCGCAAAAGGATCCA 977
Db 182 AAATTTGGATTGTGCAAGAGGCTCTCATATTAATTCATTCGCAAAAGGATCCA 241

Qy 978 CTTGACAGCTCTGAGGGCACTGTGCAAGTACAGAAATGACGCCAAGAGGATGCG 1037
Db 242 CTTGACAGCTCTGAGGGCACTGTGCAAGTACAGAAATGACGCCAAGAGGATGCG 301

Qy 1038 GCTGAGGAAGCTCCAGG-CGGCTGATGAGAAACCCGGAAGCTGAAAGATTGTGGAAG 1096
Db 302 GCTGAGGAAGCTCCAGGCGGCTGATGAGAAACCCGGAAGCTGAAAGATTGTGGAAG 361

Qy 1097 AAAAATATGAAGAAATAGCAACAGCAATGCAAGAAATGAGATCTCGAACTTTTCAGCAGC 1156
Db 362 AAAAATATGAAGAAATAGCAACAGCAATGCAAGAAATGAGATCTCGAACTTTTCAGCAGC 421

Qy 1157 TGAGTGTG-ATGTACTCCATGAAAAA 1182
Db 422 TGAGTGTGTAATGTACTTCATGAAAAA 448
```

## RESULT 14

```
US-09-764-868-241
; Sequence 241, Application US/09764868
; Patent No. US2002016871A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PT232
; CURRENT APPLICATION NUMBER: US/09/764,868
; CURRENT FILING DATE: 2001-01-17
; Prior application data removed - refer to PALM or file wrapper
; NUMBER OF SEQ ID NOS: 1510
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 241
; LENGTH: 430
; TYPE: DNA
```

```
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (286)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: SITE
; LOCATION: (342)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: SITE
; LOCATION: (347)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: SITE
; LOCATION: (371)
; OTHER INFORMATION: n equals a,t,g, or c
; US-09-764-868-241
```

```
Query Match      27.3%; Score 395.4; DB 9; Length 430;
Best Local Similarity 97.9%; Pred. No. 3.4e-102;
Matches 418; Conservative 1; Mismatches 6; Indels 2; Gaps 2;

Qy 758 CTGAGAGATATCCAAAGAACTAAATGCCATCATGGAAGCAATGTTGAACAAGATCCTT 817
Db 1 CTGAGAGATATCCAAAGAACTAAATGCCATCATGGAAGCAATGTTGAACAAGATCCTT 60

Qy 818 CATTAGACCATCTGCTATCGAAATTTTAAATAATCCCTTACCTTGTGACGACTACAGA 877
Db 61 CATTAGACCATCTGCTATCGAAATTTTAAATAATCCCTTACCTTGTGACGACTACAGA 120

Qy 878 ACCTAATGTGTAGATATTCAGAAATGACTCTGGAAGCAAAAATTTGGATTGTGCAAGG 937
Db 121 ACCTAATGTGTAGATATTCAGAAATGACTCTGGAAGCAAAAATTTGGATTGTGCAAGG 180

Qy 938 AGGTGCTCATATTAATTAATGCCATGCAAAAAGGATCCACCTGCGAGACTCTGAGGGCAC 997
Db 181 AGGTGCTCATATTAATTAATGCCATGCAAAAAGGATCCACCTGCGAGACTCTGAGGGCAC 240

Qy 998 TGTGAGAGTACAGAAATGACGCCAAGAGAGGATGCGGCTGAGAGAGCTCCAGG-CG 1056
Db 241 TGTGAGAGTACAGAAATGACGCCAAGAGAGGATGCGGCTGAGAGAGCTCCAGGCGG 300

Qy 1057 GCTGATCAGAAAGCCAGGAAGCTGAAAAAGATTGTGGAAGAAAAATATGAAGAAATAGC 1116
Db 301 GCTGATCAGAAAGCCAGGAAGCTGAAAAAGATTGTGGAAGAAAAATATGAAGAAATAGC 360

Qy 1117 AAAGCAATGCAAGAAATGAGATCTCGAACTTTTCAGCAGCTGAGTGTG-ATGTACTCCA 1175
Db 361 AAAGCAATGCAAGAAATGAGATCTCGAACTTTTCAGCAGCTGAGTGTG-ATGTACTCCA 420

Qy 1176 TGAATAA 1182
Db 421 TGAATAA 427
```

## RESULT 15

```
US-09-764-875-376
; Sequence 376, Application US/09764875
; Publication No. US20040018969A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: P0202
; CURRENT APPLICATION NUMBER: US/09/764,875
; CURRENT FILING DATE: 2001-01-17
; Prior application data removed - consult PALM or file wrapper
; NUMBER OF SEQ ID NOS: 1249
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 376
; LENGTH: 430
; TYPE: DNA
; ORGANISM: Homo sapiens
; NAME/KEY: SITE
; LOCATION: (286)
```

; OTHER INFORMATION: n equals a,t,g, or c  
; NAME/KEY: SITE  
; LOCATION: (342)  
; OTHER INFORMATION: n equals a,t,g, or c  
; NAME/KEY: SITE  
; LOCATION: (347)  
; OTHER INFORMATION: n equals a,t,g, or c  
; NAME/KEY: SITE  
; LOCATION: (371)  
; OTHER INFORMATION: n equals a,t,g, or c  
US-09-764-875-376

Query Match 27.3%; Score 395.4; DB 11; Length 430;  
Best Local Similarity 97.9%; Pred. No. 3.4e-102; Indels 2; Gaps 2;  
Matches 418; Conservative 1; Mismatches 6;  
QY 758 CTGAGAGATATCCAAAGAACTAAATGCCATCATCGAAAGCATGTTGAACAGAAATCCTT 817  
Db 1 CTGAGAGATATCCAAAGAACTAAATGCCATCATCGAAAGCATGTTGAACAGAAATCCTT 60  
QY 818 CATTAAGACCATCTGCTATCGAAATTTTAAATCCCTTACCTTGATGAGCAGCTACAGA 877  
Db 61 CATTAAGACCATCTGCTATCGAAATTTTAAATCCCTTACCTTGATGAGCAGCTACAGA 120  
QY 878 ACCTAATGTGTAGATATTCAGAAATGACTCTGGAAGACAAAAATTTGGATTGTCAGAAAG 937  
Db 121 ACCTAATGTGTAGATATTCAGAAATGACTCTGGAAGACAAAAATTTGGATTGTCAGAAAG 180  
QY 938 AGGCTGCTCATATAATTAATGCGATGCCAAAAAGGATCCACCTGCAGACTCTGAGGGCAC 997  
Db 181 AGGCTGCTCATATAATTAATGCGATGCCAAAAAGGATCCACCTGCAGACTCTGAGGGCAC 240  
QY 998 TGTGAGAAGTACAGAAAAATGACCCAAAGAGAAAGGATGCGCTGAGGAAGCTCCAGG-CG 1056  
Db 241 TGTGAGAAGTACAGAAAAATGACCCAAAGAGAAAGGATGCGCTGANGAAGCTCCAGGGCG 300  
QY 1057 GCTGATGAGAAAGCCAGGAAGCTGAAAAAGATTTGGAAGAAAAATATGAGAAAAATAGC 1116  
Db 301 GCTGATGAGAAAGCCAGGAAGCTGAAAAAGATTTGGAAGAAAAATATGAGAAAAATAGC 360  
QY 1117 AACGGAATGCAAGAAATGAGATCTCGAACTTTTCAGCAGCTGAGTGTG-ATGTACTCCA 1175  
Db 361 AACGGAATGCAAGAAATGAGATCTCGAACTTTTCAGCAGCTGAGTGTG-ATGTACTCCA 420  
QY 1176 TGA AAAA 1182  
Db 421 TGA AAAA 427

Search completed: October 3, 2004, 02:52:36  
Job time : 747 secs

---

**This Page Blank (uspto)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: October 2, 2004, 21:17:22 ; Search time 119 Seconds  
(without alignments)  
6757.343 Million cell updates/sec

Title: US-10-803-278-3

Perfect score: 1449

Sequence: 1 atgctgaattccaagagc.....actggaaaatgaaaaataa 1449

Scoring table: IDENTITY\_NUC

Gapop 10\_0 , Gapext 1.0

Searched: 682709 seqs, 277475446 residues

Total number of hits satisfying chosen parameters: 1365418

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Issued Patents NA:\*

- 1: /cgn2\_6/prodata/2/ina/5A\_COMB.seq:\*
- 2: /cgn2\_6/prodata/2/ina/5B\_COMB.seq:\*
- 3: /cgn2\_6/prodata/2/ina/6A\_COMB.seq:\*
- 4: /cgn2\_6/prodata/2/ina/6B\_COMB.seq:\*
- 5: /cgn2\_6/prodata/2/ina/PCFUS\_COMB.seq:\*
- 6: /cgn2\_6/prodata/2/ina/backfiles1.seq:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	146.2	10.1	4263	4	US-09-166-350-29 Sequence 29, Appl
2	143.2	9.9	5448	4	US-09-620-312D-246 Sequence 246, Appl
3	143.2	9.9	5532	4	US-09-620-312D-245 Sequence 245, Appl
4	131.6	9.1	2370	4	US-09-873-404-1 Sequence 1, Appl
5	104.8	7.2	1846	3	US-09-173-581-15 Sequence 15, Appl
6	104.8	7.2	1846	3	US-09-420-913-15 Sequence 15, Appl
7	89	6.1	2119	4	US-09-167-206-1 Sequence 1, Appl
8	88.6	6.1	1060	4	US-09-579-664B-2 Sequence 2, Appl
9	85	5.9	906	3	US-09-221-235-3 Sequence 3, Appl
10	85	5.9	906	3	US-09-221-928-3 Sequence 3, Appl
11	85	5.9	906	3	US-09-221-527-3 Sequence 3, Appl
12	85	5.9	906	3	US-09-221-236-3 Sequence 3, Appl
13	85	5.9	906	3	US-09-221-416-3 Sequence 3, Appl
14	85	5.9	906	3	US-09-221-245-3 Sequence 3, Appl
15	85	5.9	906	3	US-09-163-115-3 Sequence 3, Appl
16	85	5.9	906	3	US-09-221-528-3 Sequence 3, Appl
17	85	5.9	906	3	US-09-593-553-3 Sequence 3, Appl
18	85	5.9	906	3	US-09-221-237-3 Sequence 3, Appl
19	85	5.9	4137	3	US-09-221-235-1 Sequence 1, Appl
20	85	5.9	4137	3	US-09-221-928-1 Sequence 1, Appl
21	85	5.9	4137	3	US-09-221-527-1 Sequence 1, Appl
22	85	5.9	4137	3	US-09-221-236-1 Sequence 1, Appl
23	85	5.9	4137	3	US-09-221-416-1 Sequence 1, Appl
24	85	5.9	4137	3	US-09-221-245-1 Sequence 1, Appl
25	85	5.9	4137	3	US-09-163-115-1 Sequence 1, Appl
26	85	5.9	4137	3	US-09-221-528-1 Sequence 1, Appl
27	85	5.9	4137	3	US-09-593-553-1 Sequence 1, Appl

28	85	5.9	4137	3	US-09-221-237-1	Sequence 1, Appl
29	83	5.7	7218	1	US-08-232-463-14	Sequence 14, Appl
30	82.6	5.7	1251	2	US-09-211-930-2	Sequence 2, Appl
31	82.6	5.7	1251	3	US-09-340-993-2	Sequence 2, Appl
32	82.6	5.7	1251	4	US-09-468-442-2	Sequence 8, Appl
33	82.6	5.7	1353	2	US-09-211-930-8	Sequence 8, Appl
34	82.6	5.7	1353	3	US-09-340-993-8	Sequence 8, Appl
35	82.6	5.7	1353	4	US-09-468-442-8	Sequence 8, Appl
36	82.6	5.7	1542	4	US-09-345-473E-13	Sequence 13, Appl
37	82.6	5.7	3201	2	US-09-211-930-1	Sequence 1, Appl
38	82.6	5.7	3201	3	US-09-340-993-1	Sequence 1, Appl
39	82.6	5.7	3201	4	US-09-468-442-1	Sequence 1, Appl
40	82.6	5.7	3268	4	US-09-688-188B-1	Sequence 1, Appl
41	82.6	5.7	3268	4	US-09-291-417D-1	Sequence 1, Appl
42	80.2	5.5	2121	4	US-09-016-434-1142	Sequence 1142, Ap
43	75.8	5.2	1347	2	US-08-712-709-2	Sequence 2, Appl
44	75.8	5.2	1347	3	US-09-111-444-2	Sequence 2, Appl
45	75.8	5.2	1347	3	US-09-541-228-2	Sequence 2, Appl

#### ALIGNMENTS

##### RESULT 1

US-09-166-350-29

; Sequence 29, Application US/09166350A

; Patent No. 6440563

; GENERAL INFORMATION:

; APPLICANT: Scanlan, Matthew

; APPLICANT: Chen, Yao

; APPLICANT: Stockert, Elisabeth

; APPLICANT: Old, Lloyd

; APPLICANT: Jager, Elke

; APPLICANT: Knuth, Alex

; TITLE OF INVENTION: Renal Cancer Associated Antigens and

; FILE REFERENCE: L0461/7051

; CURRENT APPLICATION NUMBER: US/09/166,350A

; CURRENT FILING DATE: 1998-10-05

; EARLIER APPLICATION NUMBER: US 09/166,350

; EARLIER FILING DATE: 1998-10-05

; NUMBER OF SEQ ID NOS: 35

; SOFTWARE: RastSEQ for Windows Version 3.0

; SEQ ID NO 29

; LENGTH: 4263

; TYPE: DNA

; ORGANISM: Homo sapiens

US-09-166-350-29

Query Match 10.1%; Score 146.2; DB 4; Length 4263;

Best Local Similarity 52.5%; Pred. No. 3.7e-33;

Matches 372; Conservative 0; Mismatches 328; Indels 9; Gaps 2;

QY	173	AGGTACTTAAGGAATATCTGTTGGAGAACTAAATCCAAATGAAACTGTACAGCCCAATT	232
Db	664	ATGTCATCAAGGAATTAACATCTCAAGAATGTCGTATAAGAAAGCAAGAAATCAAGGA	723
QY	233	TGGAGCCCAACTCTCTCCAGCTGGACCCAGCCATGTCAAGTTCATGCAAGTT	292
Db	724	GAGAAATGCTGTATGTCGCAACCAATGAAGCATCCAAATATTGTCCAATATAAGAAATCAT	783
QY	293	TTGTGGAGCAAGATAAATTTCTGCATTATCACGGAGTACTGTGAGGGCCGAGATCTGAGCG	352
Db	784	TTGAAGAAATGGCTCTCTCTACATAGTAAAGAGGAGGATTTGTTTA	843
QY	353	ATAAAATTCAGGAATATAACACAGCTGGAAATCTTTCCAGAAATCAATATAAGAAAT	412
Db	844	AACGAATAAATGCTCAGAAAGGCGCTC-----TGTTTCAAGAGACACAGATTGGAAT	897
QY	413	GGTTTATCCAGCTGCTGCTGGAGTTTGACTACATGATGAGAGGAGATCTTCTCATCGAG	472
Db	898	GGTTTGTGCAGATATGTTTGGCTCTGAAGCATGTACATGATAGAAAATTTCTTCACCGAG	957

473	QY	AC	TTAAAGTCAAAAGATGTA	TTTCT--GAAAAATAATCTCTTAAATTTGGAGATTTTG	529
958	Db	AC	TAAAGTCACAGAACATATTTCTAACCAAAAGATGGACAGTGCAGCTTGGAGATTTTG	1017	
530	QY	GAGTTTCTCGACTTC	TAAATCGGATCTGTGACTGCGCCACAACTTTAACTGGAACTCCCC	589	
1018	Db	GAA	TTGCTCGAGTTC	TAAATAGTACTGTAGAGTGGCTCGAACTTGCATAGGCACCTCCAT	1077
590	QY	ATTATATGAGTCTGAGGCTCTGAAAAACCAAGGCTATGACACAAAAGTCGGACATCTG	T	649	
1078	Db	ACTACTGTGCACCTGAAATCTGTGAAAAACAAGCCTTATAACATATAAGTGCATTTGG	1137		
650	QY	CAC	TGCATGCATTTTGTATAGATGTGTGCAAGAAATCATGCATTCGTGGCTCCAAAT	709	
1138	Db	CTTTGGGCTGTGCTCTTATAGTTGTGTACACTTAAACATGCA	TTTGAAGCTGGAACA	1197	
710	QY	TCTTATCCATTGTTTAAAAATTTGTAAGGTGACACACTTCTCTCCCTGAGAGATATC	769		
1198	Db	TGAAAAACCTGGTACTGGAAGATAATCTCCGGATCCTTCTCCAGTGTCTCCACATTA	CT	1257	
770	QY	CAAAAGCACTAAATGCCATCATGGAAGCATGTTGAACCAAGAAATCCTTCATTAAGACCAT	829		
1258	Db	CCTATGATCTCCGAGCTGCTGCTCAGTTATTTAAAGAAATCCTAGGATAGACCAT	1317		
830	QY	CTGCTATCGAAATTTTAAAAATCCCTTACCTTGATGAGCAGCTACAGAA	878		
1318	Db	CAGTCAACTCCATATTGGAGAAAGTTTTATAGCTTAACGAATTCGAAA	1366		

## RESULT 2

RESULT 2  
US-09-620-312D-246  
US-09-620-312D-246, Application US/09620312D  
; Patent No. 6589662  
; GENERAL INFORMATION:  
; APPLICANT: Tang, Y. Tom  
; APPLICANT: Liu, Chenghua  
; APPLICANT: Asundi, Vinod  
; APPLICANT: Zhang, Jie  
; APPLICANT: Ren, Feiyang  
; APPLICANT: Chen, Rui-hong  
; APPLICANT: Zhao, Qing A.  
; APPLICANT: Wehrman, Tom  
; APPLICANT: Xue, Aidong J.  
; APPLICANT: Yang, Yonghong  
; APPLICANT: Wang, Jian-Rui  
; APPLICANT: Zhou, Ping  
; APPLICANT: Ma, Yunding  
; APPLICANT: Wang, Dunrui  
; APPLICANT: Wang, Zhiwei  
; APPLICANT: John Tillinghast

Query Match 9.9%; Score 143.2; DB 4;  
Best Local Similarity 52.3%; Pred. No. 3.3e-32;  
Length 5448;

Matches	369;	Conservative	0;	Mismatches	328;	Indels	9;	Gaps	2;
Qy	148	AAAGCCAAACGAGGAGAGGAATTAAGGTACTTAAGGAAATATCTGTTGGAGAACTAAAT	207						
Db	597	AAATCTACAGAAAGATGGCAGACAGTAGTATGTTATCAAGGAAATTAACATCTCAAGAAATGTCC	656						
Qy	208	CCAAATGAACTGTACAGGCCCAATTTGGAGCCCACTCCTCTCCAAGCTGGACCAACCCA	267						
Db	657	AGTAAAGAAGAAGAAATCAAGGAGAGAGTTGCAAGTATTTGGCAAAACATGAAGCATCCA	716						
Qy	268	GCCATTGTCAAGTTCATGCAAGATTTTGTGGAGCAAGATAATTTCTGCATTATCACGGAG	327						
Db	717	ATAATTGTCAGTATAGAGAAATCATTTGAAGAAATGGCTCTCTCTACATAGTAATGGAT	776						
Qy	328	TACTGTGAGGCGGAGATCTGAGACGATAAAATTCAGGAATATAAACAAGCTGGAANAATC	387						
Db	777	TACTGTGAGGAGGGGATCTGTTTAAGCGAAATAAATGCTCAGAAAGCGGTT-----TG	830						
Qy	388	TTTCCAGAAAATCAAAATAATAGAATGGTTTATCCAGCTGCTGTGCGGAGTTCACTACATG	447						
Db	831	TTTCAGAGGATCAGATTTTGGACTGTTTGTACAGATATGTTTGGCCCTGAAACATGTA	890						
Qy	448	CATGAGAGAGGATCTCATCGAGACTTAAGTCAAGAAATGTTATTTCT---GAAAAAT	504						
Db	891	CATGATAGAAAATTTCTTCATCGAGACATAAATCTCAGAACATATTTTTTAACATAAGAT	950						
Qy	505	AATCTCTTTAAAATTTGGAGTTTCTCGACTTCTAATGGGATCTGTGACCTTG	564						
Db	951	GGAA CAGTACAACTTGGAGATTTTGGAAATGCTAGAGTTCTTAATAGTACTGTAGAGCTG	1010						
Qy	565	GCACAACCTTAACCTGGAACTCCCACTAATAGATGCTCGAGGCTCGAAACCAAGGC	624						
Db	1011	GCTCGAACTTGCATAGGGACCCACTACTTGTCACTGAAATCTGTGAAACCAACCT	1070						
Qy	625	TATGACACAAAGTCGAGACATCTGGTCACGTGGCATGCATTTGTATGAGATGTGCTGCATG	684						
Db	1071	TACAATAATAAAGTGACATTTTGGGCTCTGGGGTGTCTCTTTATGAGCTGTGTACATT	1130						
Qy	685	AATCATGCATTCGCTGGCTCCAATTTCTTATCATTGTTTTAAAAATTTGTTGAAGGTGAC	744						
Db	1131	AAACATGCTTTTGAAGCTGGCAGTATGAAAAACCTGGTACTGAAGATAATATCTCGATCT	1190						
Qy	745	ACACCTTCTCTCCCTGAGAGATATCCAAAGAACATAAATGCCATCATGGAAGCATGTTG	804						
Db	1191	TTTCCACCTGTGCTTTGTCATTATTCCTAATGATCTCCGAGTTTGGTGTCTCAGTTATT	1250						
Qy	805	AACAAGAATCTCTTCAATTAAGACCATCTGCTATCGAAATTTTAAAA	850						
Db	1251	AAAAGAAATCTTAGGATAGACCAATCAGTCAACTCCATATTGGAGA	1296						

### RESULT 3

```

US-09-620-312D-245
; Sequence 245, Application US/09620312D
; Patent No. 6569662
;
GENERAL INFORMATION:
;
; APPLICANT: Wang, Y. Tom
; APPLICANT: Liu, Chenchua
; APPLICANT: Asundi, Vinod
; APPLICANT: Zhang, Jie
; APPLICANT: Ren, Feiyan
; APPLICANT: Chen, Rui-hong
; APPLICANT: Zhao, Qing A.
; APPLICANT: Wehrman, Tom
; APPLICANT: Xue, Aidong J.
; APPLICANT: Yang, Yongchong
; APPLICANT: Wang, Jian-Rui
; APPLICANT: Zhou, Ping
; APPLICANT: Ma, Yunqing
; APPLICANT: Wang, Dunrui
; APPLICANT: Wang, Zhiwei
; APPLICANT: John Tillinghast
; APPLICANT: Drmanac, Radoje T.

```

; TITLE OF INVENTION: No. 6569662el Nucleic Acids and  
; FILE REFERENCE: Polypeptides  
; CURRENT APPLICATION NUMBER: US/09/620,312D  
; PRIOR FILING DATE: 2000-07-19  
; PRIOR APPLICATION NUMBER: 09/552,317  
; PRIOR FILING DATE: 2000-04-25  
; PRIOR APPLICATION NUMBER: 09/488,725  
; PRIOR FILING DATE: 2000-01-21  
; NUMBER OF SEQ ID NOS: 1105  
; SOFTWARE: pt\_FL\_genes Version 1.0  
; SEQ ID NO 245  
; LENGTH: 5532  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: (534)..(4262)  
US-09-620-312D-245

Query Match 9.9%; Score 143.2; DB 4; Length 5532;  
Best Local Similarity 52.3%; Pred. No. 3.4e-32;  
Matches 369; Conservative 0; Mismatches 328; Indels 9; Gaps 2;  
QY 148 AAGGCCAAACGAGGAGGAGGAATTAAGGTACTTAAGGAAATATCTGTTGGAGAACTAAAT 207  
DB 597 AAATCTACAGAAGATGGCAGACAGTATGTTATCAAGGAATTAACATCTCAAGATGTCC 656  
QY 208 CCAATGAATCTGTACAGGCCAATTTGGAGCCCACTCTCCAGCTGGACCAACCA 267  
DB 657 AGTAAAGAAGAAGAAGATCAAGGAGAGAAATGTCAGTATTCGCAACATGAAGCATCCA 716  
QY 268 GCCATTGTCAAGTTCATCAAGATTTTGGAGCAAGATAAATTTCTGCAATATCACGGAG 327  
DB 717 AATATGTCCAGTATAGAGAATCAATTTGAAGAAATGGCTCTCTACATAGTAATGAT 776  
QY 328 TACTGTGAGGGCCGAGATCTGACGACGATAAAATTCAGGAATATAAACAAGCTGGAAAAATC 387  
DB 777 TACTGTGAGGGAGGGGATCTGTTAAGCGAATAAATTCGTCAGAAAGGGGTTT-----TG 830  
QY 388 TTTCAGAAATACAAATATAGATAGTGGTTTATCCAGCTGCTGCTGGAGTTTACATG 447  
DB 831 TTTCAGAGAGATCAGATTTTGGACTGGTTTACAGATATGTTGGCCCTGAAACATGTA 890  
QY 448 CATGAGAGGAGATCTTTCATCGAGACTTAAAGTCAAAGAATGTATTTCT---GAAAAAT 504  
DB 891 CATGATAGAAAAATTTCTTCATCGAGACATTAATATCTCAGAACATATTTTAACTAAAGAT 950  
QY 505 AATCTCCTTAAATTCGAGATTTTGGAGTTTCTGACCTTCTAATGGGATCTGTGACCTG 564  
DB 951 GGAACAGTACAACTTGGAGATTTTGGAAATTTGCTAGAGTTCTTAAATAGTACTGTAGAGCTG 1010  
QY 565 GCCACAACTTTAACTGGAACTCCCAATATATGAGTCTGAGGCTCTGAAACACCAAGGC 624  
DB 1011 GCTCGAATTCGATAGGAGCCCACTACTTGTACCTGAAATCTGTGAAACCAACCT 1070  
QY 625 TATGACAAAAGTCGGACATCTGGTCACTGGCAATGCAATTTGTATGAGATGTGCTGCAATG 684  
DB 1071 TACAATAATAAAGTACATTTGGGCTCTGGGGTGTGCTCTTATGAGCTGTGACACTT 1130  
QY 685 AATCATGCAATCTGCTGGCTCCAAATTTCTTATCCATTTTAAATTTTAAATTTTGAAGTGC 744  
DB 1131 AAACATGCTTTGAAGCTGGCAGTATGAAGAACCTGTACTGAGATATATCTGATCT 1190  
QY 745 ACACTTCTCTCCCTGAGAGATATCAAAAGAACTAAATGCCATCATGGAAGCAATGTTG 804  
DB 1191 TTTCCACCTGTGCTTTTGGCATTTATCTATGATCTCCGAGTTTGGTGTCTCAGTATTT 1250  
QY 805 AACAGAAATCTTCATTAAGACCATCTGCTATCGAAATTTTAAAA 850  
DB 1251 AAAAGAAATCTTAGGATAGACCATCAGTCAACTCCATATTGGAGA 1296

RESULT 4  
US-09-873-404-1  
; Sequence 1, Application US/09873404  
; Patent No. 6500656  
; GENERAL INFORMATION:  
; APPLICANT: WEBSTER, Marion et al  
; TITLE OF INVENTION: ISOLATED HUMAN KINASE PROTEINS, NUCLEIC  
; TITLE OF INVENTION: ACID MOLECULES ENCODING HUMAN KINASE PROTEINS, AND USES  
; TITLE OF INVENTION: THEREOF  
; FILE REFERENCE: CL001212-CIP  
; CURRENT APPLICATION NUMBER: US/09/873,404  
; CURRENT FILING DATE: 2001-06-05  
; NUMBER OF SEQ ID NOS: 4  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 1  
; LENGTH: 2370  
; TYPE: DNA  
; ORGANISM: Human  
US-09-873-404-1

Query Match 9.1%; Score 131.6; DB 4; Length 2370;  
Best Local Similarity 51.3%; Pred. No. 5.9e-29;  
Matches 364; Conservative 0; Mismatches 334; Indels 12; Gaps 2;  
QY 175 GTACTTAAGGAAATATCTGTTGGAGAACTAAATCCAAATGAAACTGACAGGCCAATTTG 234  
DB 102 GTCATAAAGAGATCAATTTTGAAGAGATGCCCATACAGAAAAAGAGCTTCAAGAAA 161  
QY 235 GAAGCCCAACCTCTCCAGCTGRCACCCAGCCCACTTGTCAAGTTCATGCCAAGTTT 294  
DB 162 GAAGTGAATCTTCTGGAAAGATGAAACATCCCAACATTTAGCCTTCTTCAATTCATTT 221  
QY 295 GTGAGCAAGATTAATTTCTGCATTATCACGAGAGTACTGTGAGGCCGAGATCTGAGCAT 354  
DB 222 CAAGAGATGGCAGCTGTTTATGTAATGGAATATTGTGATGGAGGGATCTCATGAAA 281  
QY 355 AAAATTGAGGAATATAACAGCTGGAAAAATCTTCCAGAAATCAATTAATAGATGG 414  
DB 282 AGGATCA-----ATAGACAAACGGGGTGTGTATTTAGTGAAGATCAGATCTCGGTGG 335  
QY 415 TTTATCAGCTGCTGCTGGAGTTGACTACATGATGAGAGGAGGATCTTCACTCGAGAC 474  
DB 336 TTGTACAGATTTCTCTAGGACTAAACATATTCATGACAGGAGATATTACACAGGAC 395  
QY 475 TTAAGTCAAAGATGTATTTCTGAAAAATAA-----TCTCCTTAAATTTGGAGATTT 528  
DB 396 ATAAAGCTCAGAACATTTTCTTAGCAAGAACGGAATGGTGGCAAGCTTGGGGACTTT 455  
QY 529 GGAGTTTCTCGACTTCTAAATGGGATCTGTGACTGGCCCACTTTTAACTGGAACCTCC 588  
DB 456 GGTATAGCAGAGTCTCTGAATAATTCATGGAATCTGCTGAACTTGTATTTGGAACACT 515  
QY 589 CATTATATGAGTCTCTGAGGCTCTGAAACACCAAGGCTATGACACAAAGTCTGGACATCTGG 648  
DB 516 TACTACCTGTCCCCAGAGATCTGTGAGAATAAACCCCTACAAACAATAAAACGGATATTGG 575  
QY 649 TCATCGCATGTCATTTTGTATGAGATGTGCTGATGATGATCATGATTCGCTGGCTCCAAT 708  
DB 576 TCTCTTGGCTGTGCTTATATGAGCTCTGCACACTTAAACATCTTTTGGGGTAAACAC 635  
QY 709 TCTTATTCATTTTAAATAATTTGAAAGGTGACACACTTCTCTCCTCTGAGAGATAT 768  
DB 636 TTACAGCAGCTGGTTCTGAAGATTTGTCAAGCACATTTTCCCAATATCTCCGGGGTTT 695  
QY 769 CCAAAAGAACTAAATGCGCATCATGGAAGCATGTTGAAACAGAAATCTTTCATTAAGACCA 828  
DB 696 TCTCGTGAAGTCCATTCCTTGATATCTCAGCTCTTTCAGTATCTCTCTCGAGACCGACCA 755  
QY 829 TCTGCTATCGAAATTTTAAAAATCCCTTACTTCTGATGAGCAGCTTACAGAA 878  
DB 756 TCCATAAATTCCTATTTTGAAGAGCCCTTTTAGAGATCTTATTCCCAA 805



RESULT 5  
US-09-173-581-15  
; Sequence 15, Application US/09173581A  
; Patent No. 6013455  
; GENERAL INFORMATION:  
; APPLICANT: Bandman, Olga  
; APPLICANT: Tang, Y. Tom  
; APPLICANT: Hillman, Jennifer L.  
; APPLICANT: Yue, Henry  
; APPLICANT: Guegler, Karl J.  
; APPLICANT: Corley, Neil C.  
; APPLICANT: Gorgone, Gina  
; APPLICANT: Azimzai, Yalda  
; APPLICANT: Lu, Aina  
; TITLE OF INVENTION: Protein Kinase Homologs  
; FILE REFERENCE: PF-0614 US  
; CURRENT APPLICATION NUMBER: US/09/173,581A  
; CURRENT FILING DATE: 1998-10-15  
; EARLIER FILING DATE: 1998-10-15  
; NUMBER OF SEQ ID NOS: 18  
; SOFTWARE: PERL Program  
; SEQ ID NO 15  
; LENGTH: 1846  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
; FEATURE: -  
; OTHER INFORMATION: 1567782  
US-09-173-581-15

Query Match 7.2%; Score 104.8; DB 3; Length 1846;  
Best Local Similarity 53.5%; Pred. No. 5.1e-21;  
Matches 242; Conservative 0; Mismatches 207; Indels 3; Gaps 1;  
QY 402 AATAATAGATGTTTATCCAGCTGCTGGAGTTGACTACATGCTGAGAGGAT 461  
DB 369 AATTTGGAGTGGTTGTACAGATATGTTGGCCCTGAAACATGTACATAGAAAAAT 428  
QY 462 ACTTCATCGAGACTTAAAGTCAAAGATGTATTTCTGAAAAATAATCTCC---TTAAAT 518  
DB 429 TCTTCATCGAGACATTAATCTCAGAACATATTTTAACTAAGATGGAACAGTACAAC 488  
QY 519 TGGAGATTTGGAGTTTCTCGACTTCTAATGGATCTGTGACCTGGCCACAACTTTAAC 578  
DB 489 TGGAGATTTTGGAAATTCGTAGAGTTCTTAATAGTACTGTAGAGCTGGCTCGAACTTGCAT 548  
QY 579 TGGAACTCCCAATATATAGTCTCGAGGCTCTGAAACACCAAGGCTATGACACAAAGTC 638  
DB 549 AGGACCCCATACTACTTGTACCTGAAATCTGTGAAACAAACCTTACAAATAAAG 608  
QY 639 GGACATCTGGTCTACCTGGCATGCAATTTGTATGAGATGTCTGATGAATCATGCAATTCGC 698  
DB 609 TGACATTTGGGCTCTGGGCTGTCTCTTTATGAGCTGTGTACACTTAAACATGCTTTGA 668  
QY 699 TGGCTCCAAATTTCTATCCATTTGTTTAAATAATTTGTAAGGTGACACACCTTCTCTCC 758  
DB 669 AGCTGGCAGTATGAAACAACTGTGATCTGAAGATAATATCTGGATCTTTCCACCTGTCTC 728  
QY 759 TGAGAGATATCCAAAAGAACTAAATCCCATCATGGAAGCATGTTGAACAAGAACTCTTC 818  
DB 729 TTTGCATTTATCCATGATCTCCGAGTTTGGTGTCTCAGTTATTTAAAGAAATCCTAG 788  
QY 819 ATTAAGACCATCTGCTATCGAAATTTTAAAA 850  
DB 789 GGATAGACCATCATGCTCAACTCCCATATTGGAGA 820

RESULT 6  
US-09-420-915-15  
; Sequence 15, Application US/09420915  
; Patent No. 6264947  
; GENERAL INFORMATION:  
; APPLICANT: Bandman, Olga  
; APPLICANT: Tang, Y. Tom  
; APPLICANT: Hillman, Jennifer L.

APPLICANT: Yue, Henry  
APPLICANT: Guegler, Karl J.  
APPLICANT: Corley, Neil C.  
APPLICANT: Gorgone, Gina  
APPLICANT: Azimzai, Yalda  
APPLICANT: Lu, Aina  
TITLE OF INVENTION: Protein Kinase Homologs  
FILE REFERENCE: PF-0614 US  
CURRENT APPLICATION NUMBER: US/09/420,915  
CURRENT FILING DATE: 1999-10-20  
EARLIER APPLICATION NUMBER: US 09/173,581  
EARLIER FILING DATE: 1998-10-15  
NUMBER OF SEQ ID NOS: 18  
SOFTWARE: PERL Program  
SEQ ID NO 15  
LENGTH: 1846  
TYPE: DNA  
ORGANISM: Homo sapiens  
FEATURE: -  
OTHER INFORMATION: 1567782  
US-09-420-915-15

Query Match 7.2%; Score 104.8; DB 3; Length 1846;  
Best Local Similarity 53.5%; Pred. No. 5.1e-21;  
Matches 242; Conservative 0; Mismatches 207; Indels 3; Gaps 1;  
QY 402 AATAATAGATGTTTATCCAGCTGCTGGAGTTGACTACATGCTGAGAGGAT 461  
DB 369 AATTTGGAGTGGTTGTACAGATATGTTGGCCCTGAAACATGTACATAGAAAAAT 428  
QY 462 ACTTCATCGAGACTTAAAGTCAAAGATGTATTTCTGAAAAATAATCTCC---TTAAAT 518  
DB 429 TCTTCATCGAGACATTAATCTCAGAACATATTTTAACTAAGATGGAACAGTACAAC 488  
QY 519 TGGAGATTTGGAGTTTCTCGACTTCTAATGGATCTGTGACCTGGCCACAACTTTAAC 578  
DB 489 TGGAGATTTTGGAAATTCGTAGAGTTCTTAATAGTACTGTAGAGCTGGCTCGAACTTGCAT 548  
QY 579 TGGAACTCCCAATATATAGTCTCGAGGCTCTGAAACACCAAGGCTATGACACAAAGTC 638  
DB 549 AGGACCCCATACTACTTGTACCTGAAATCTGTGAAACAAACCTTACAAATAAAG 608  
QY 639 GGACATCTGGTCTACCTGGCATGCAATTTGTATGAGATGTCTGATGAATCATGCAATTCGC 698  
DB 609 TGACATTTGGGCTCTGGGCTGTCTCTTTATGAGCTGTGTACACTTAAACATGCTTTGA 668  
QY 699 TGGCTCCAAATTTCTATCCATTTGTTTAAATAATTTGTAAGGTGACACACCTTCTCTCC 758  
DB 669 AGCTGGCAGTATGAAACAACTGTGATCTGAAGATAATATCTGGATCTTTCCACCTGTCTC 728  
QY 759 TGAGAGATATCCAAAAGAACTAAATCCCATCATGGAAGCATGTTGAACAAGAACTCTTC 818  
DB 729 TTTGCATTTATCCATGATCTCCGAGTTTGGTGTCTCAGTTATTTAAAGAAATCCTAG 788  
QY 819 ATTAAGACCATCTGCTATCGAAATTTTAAAA 850  
DB 789 GGATAGACCATCATGCTCAACTCCCATATTGGAGA 820

RESULT 7  
US-09-167-206-1  
; Sequence 1, Application US/09167206A  
; Patent No. 6476193  
; GENERAL INFORMATION:  
; APPLICANT: Nandabalan, Krishnan  
; APPLICANT: Schulz, Vincent P.  
; APPLICANT: Yang, MeiJa  
TITLE OF INVENTION: NIK1 PROTEIN AND NIK1 PROTEIN COMPLEXES  
FILE REFERENCE: 15966-521 NIK1 protein complexes  
CURRENT APPLICATION NUMBER: US/09/167,206A  
CURRENT FILING DATE: 1998-10-06  
NUMBER OF SEQ ID NOS: 26  
SOFTWARE: Patent In Ver. 2.0

```
; SEQ ID NO 1
; LENGTH: 2119
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (135)..(1469)
US-09-167-206-1

Query Match
Best Local Similarity 6.1%; Score 89; DB 4; Length 2119;
Matches 216; Conservative 0; Mismatches 190; Indels 3; Gaps 1;

QY 461 TACTTCATCGAGACTTAAAGTCAAGAAATGATTTCTGNA---AAATATCTCCTTAAAA 517
D 544 TATTGATCGGATCTTAAACAGGCAATGTTTCTCGATGCGCAAGCAAAACGTCAGC 603
QY 518 TTGGAGATTTTGGAGTTTCTCCACTTCTAAATGGGATCCTGTGACCTGGGCCCAAACTTTAA 577
D 604 TTGGAGATTTTGGCTAGCTAGTAATTAACCATGACACAGAGTTTTCNAAAAAATTTG 663
QY 578 CTGGAACCTCCCATATATAGTCTGAGGCTCTGAACACCAAGGCTATGACAAAGT 637
D 664 TTGGCACACCTTATTACATGTCTCTGAAACAAATGAATCGCATGTCTACAAATGAGAAAT 723
QY 638 CGGACATCTGGTCACTGGCATGATTTTCTATGAGATGTGCTGCAATGATCATGCTCG 697
D 724 CAGATATCTGGTCACTGGCTGCTGCTGATGAGTTAGTTATGCTCATTAATGCTCCTCA 783
QY 698 CTGGCTCCAAATTTCTATCCATTTGTTTAAATAATTTTGAAGGTGACACACCTTCTCC 757
D 784 CAGCTTTTAGCCAGAAAGAACTCGCTGGGAAATCAGAGAGGCAAAATTCAGGCGAATTC 843
QY 758 CTGAGAGATATCCAAAGAACTAAATGCCATCATGGAAGCATGTGAACAGATCTCT 817
D 844 CATACCGTTACTCTGATGAATTAATGAATTAATTAACGAGGATGTTAAACTTAAAGGAT 903
QY 818 CATTAAAGACCATCTGCTATCGAAATTTTAAATAATCCCTTACCTTGATGA 866
D 904 ACCATCGACCTTCTGTTGAAGAAATTTCTTGAGAACCCITTAATAGCAGA 952

RESULT 8
US-09-579-664B-2
; Sequence 2, Application US/09579664B
; Patent No. 6514719
; GENERAL INFORMATION:
; APPLICANT: Immunex Corporation
; APPLICANT: Bird, Timothy A.
; APPLICANT: Virca, G. Duke
; APPLICANT: Martin, Unja
; APPLICANT: Anderson, Dirk M.
; TITLE OF INVENTION: NOVEL MURINE AND HUMAN KINASES
; FILE REFERENCE: 2923-A
; CURRENT APPLICATION NUMBER: US/09/579,664B
; CURRENT FILING DATE: 2000-05-26
; NUMBER OF SEQ ID NOS: 36
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 2
; LENGTH: 1060
; TYPE: DNA
; ORGANISM: Mus musculus
US-09-579-664B-2

Query Match
Best Local Similarity 6.1%; Score 88.6; DB 4; Length 1060;
Matches 269; Conservative 0; Mismatches 279; Indels 3; Gaps 1;

QY 160 GGAGAGAAATTAAGTACTTAAAGAAATATCTGTTGGAGAACTAAATCCAAATGAAGT 219
D 283 GGAGTCCCGTAGCGTTAAAGAAAGTACAGATATTGATTTAATGATGCGCAAGCAGT 342
QY 220 GTACAGGCCAAATTTGAAGCCCACTCTCTCAAGCTGGACCCAGCCATTTGTCAAG 279
```

```
Db 343 GCTGATTGTATCAAGAAATAGACCTCTCTTAAGCAACTCAACCATCCAAATGTAATTA 402
QY 280 TTCCATCAAGATTTTGTGGAGCAAGATAATTTCTGCAATTAATCAGGAGTACTGTGAGG 339
D 403 TACTATCATCATTAATTTGAGGATAATGAGCTGAACATAGTTTGGAGTTAGCAGATGCT 462
QY 340 CGAGATCTGGACGATAAATTCAGGAATATAACAGCTGGAAAAATCTTTCCAGAAAT 399
D 463 GGTGACCTCTCCAGAAATGATAAAGCACTTTAAGAAAAAAGAGGCTAAATCCCTGAGAG 522
QY 400 CAAATAATAGAAATGTTTATCCAGCTGCTCTGGAGTTGACTACATGCAATGAGAGGAGG 459
D 523 ACCGTTTGGAAATATCTTCTGCTGCTGAGTGCAGTGCACCTGACCATTCCTCGAAGA 582
QY 460 ATACTTCATCGAGACTTAAAGTCAAGAAATGATTTCTGAAA---AATAATCTCCTTTAA 516
D 583 GTCATGCACAGAGATATAAAACCAAGCTAATGTGTTCATTACAGCCACTGGGGTAGTAAA 642
QY 517 ATTGGAGATTTTGGAGTTTCTCGACTTCTAATGGGATCCTGTGACCTGGGCCACAACTTT 576
D 643 CTGGAGACCTTGGGCTTGTGGTCTTTCAGCTCCAAACCAAGCTGACATCTTTTA 702
QY 577 ACTGGAACCTCCCATTAATATGAGTCTTGAGGCTCTGAAACCAAGGCTATGACAAAG 636
D 703 GTGGGTACACCTTACTACTATGTCTCCAGAGAGAAATACATGAAAAATGGATACAACTT 762
QY 637 TCGAGACATCTGCTCACTGGCATGCACTTTTGTATGAGATGTGCTGCATGATCATGCTTC 696
D 763 TCTGACATCTGGTCTCTTGGCTGTCTGCTATATGAGATGGCTGCACTGCAGAGTCTTTC 822
QY 697 GCTGGCTCCAA 707
D 823 TACGGCGACAA 833

RESULT 9
US-09-221-235-3
; Sequence 3, Application US/09221235
; Patent No. 6043040
; GENERAL INFORMATION:
; APPLICANT: Acton, Susan
; TITLE OF INVENTION: NOVEL CSAPK-1 NUCLEIC ACID MOLECULES AND USES THEREFOR
; FILE REFERENCE: NNI-050
; CURRENT APPLICATION NUMBER: US/09/221,235
; CURRENT FILING DATE: 1998-12-28
; EARLIER APPLICATION NUMBER: 09/163,115
; EARLIER FILING DATE:
; NUMBER OF SEQ ID NOS: 15
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 3
; LENGTH: 906
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(906)
US-09-221-235-3

Query Match
Best Local Similarity 5.9%; Score 85; DB 3; Length 906;
Matches 254; Conservative 0; Mismatches 260; Indels 3; Gaps 1;

QY 194 TTGGAGAACTAAATCCAAATGAAACTGTACAGGCCAAATTTGGAGCCCACTCTCTTCCA 253
D 203 TTGATTTAATGGATGCCAAAGACGCTGCTGATTCATCAAGAAATAGATCTTCTTAAAGC 262
QY 254 AGCTGGACCCACCCAGCCATTGTCAAGTTCCATGCAAGTTTGTGGACCAAGATAATTTCT 313
D 263 AACTCAACCATCCAAATGTAATAAATATTATGATCATCTTTCATGAGATTAAGAACTAA 322
QY 314 GCATTATACCGAGTACTGTGAGGCGCGAGATCTGGACGATGGAATAAATTCAGAAATATAAC 373
```

Db 323 ACATAGTTTGGAACTAGCAGATGCTGGCGACCTATCCAGATGATCAAGCATTTTAAAGA 382  
Qy 374 AAGCTGGAATAATCTTTCCAGAAATCAATAATAGAAATGGTTTATCCAGCTGCTGCTGG 433  
Db 383 AGCAAAAGAGGCTAAATCTCTGAAAGAACTGTTTGGAAATATTTTGTTCAGCTTTCAGTG 442  
Qy 434 GAGTTGACTACATGCGATGAGAGAGAGATCTTCATCGAGACTTAAAGTCAAGAAATGTAT 493  
Db 443 CATTTGAAACACATGCAATCTCGAAGAGTCAATGATAGATATAAAACCAAGTAAATGTGT 502  
Qy 494 TTCTGAAA---AATAATCTCTTTAAATTTGGAGATTTTGGAGTTTCTCGACCTTCTAATGG 550  
Db 503 TCATTACAGCCACTGGGGTGTAAACTTGGAGATCTTGGCGCTTGGCGGTTTTTTCAGCT 562  
Qy 551 GATCCTGTGACCTGGCCACAACTTTAACTGGAATCTCCCATTTATATGAGTCCTGAGGCTC 610  
Db 563 CAAAACCAAGGCTATGACACAAAGTGGACATCTGGTCACTGGCATCATTTTGTATG 670  
Qy 611 TGAAACACCAAGGCTATGACACAAAGTGGACATCTGGTCACTGGCATCATTTTGTATG 670  
Db 623 TACATGAAATGGATGATCACTTCAAACTGACATCTGGTCTTGGCTGTCTACTATATG 682  
Qy 671 AGATGCTGCTGCAATGATCATGCTTGGCTGGCTCCAA 707  
Db 683 AGATGGCTGCAATGATCAAAAGTCTTCTATGGTGACAA 719

RESULT 10  
US-09-221-928-3  
; Sequence 3, Application US/09221928  
; Patent No. 6121030  
; GENERAL INFORMATION:  
; APPLICANT: Acton, Susan  
; TITLE OF INVENTION: NOVEL CSAPK-1 NUCLEIC ACID MOLECULES AND USES THEREFOR  
; FILE REFERENCE: MNI-050  
; CURRENT APPLICATION NUMBER: US/09/221,928  
; EARLIER FILING DATE: 1998-12-28  
; EARLIER FILING DATE:  
; NUMBER OF SEQ ID NOS: 15  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 3  
; LENGTH: 906  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: (1)..(906)  
US-09-221-928-3

Query Match 5.9%; Score 85; DB 3; Length 906;  
Best Local Similarity 49.1%; Pred. No. 2.7e-15;  
Matches 254; Conservative 0; Mismatches 260; Indels 3; Gaps 1;  
Qy 194 TTGGAGAACTAAATCCAAATGAACTGTACAGGCCAATTTGGAGCCCACTCTCTCCA 253  
Db 203 TTGATTTAATGGATGCCAAGCAGCTGCTGATTCATCAAGAAATAGATCTTCTTAAGC 262  
Qy 254 AGCTGGACCAACCCAGCCATTTGTCAGTTCATGCAAGTTTGGAGCAAGATAATTTCT 313  
Db 263 AACTCAACCATCCAAATGTAATAAATATTTATGATCAATCAAGATATGAATGA 322  
Qy 314 GATTTACACGAGTACTGTGAGGCGGAGATCTGACGATATAAATTTTGGAGCAAGATAATTTCT 373  
Db 323 AACTCAACCATCCAAATGTAATAAATATTTATGATCAATCAAGATATGAATGA 382  
Qy 374 AAGCTGGAATAATCTTTCCAGAAATCAATAATAGAAATGGTTTATCCAGCTGCTGCTGG 433  
Db 383 AGCAAAAGAGGCTAAATCTCTGAAAGAACTGTTTGGAAATATTTTGTTCAGCTTTCAGTG 442  
Qy 494 TTCTGAAA---AATAATCTCTTTAAATTTGGAGATTTTGGAGTTTCTCGACTTCTAATGG 550  
Db 503 TCATTACAGCCACTGGGGTGTAAACTTGGAGATCTTGGCGCTTGGCGGTTTTTTCAGCT 562  
Qy 551 GATCCTGTGACCTGGCCACAACTTTAACTGGAATCTCCCATTTATATGAGTCTGAGGCTC 610  
Db 563 CAAAACCAAGCTGCAATCTTTAGTTGGTACGCTTATTACATGTCTTCCAGAGAA 622

Qy 494 TTCTGAAA---AATAATCTCTTTAAATTTGGAGATTTTGGAGTTTCTCGACTTCTAATGG 550  
Db 503 TCATTACAGCCACTGGGGTGTAAACTTGGAGATCTTGGCGCTTGGCGGTTTTTTCAGCT 562  
Qy 551 GATCCTGTGACCTGGCCACAACTTTAACTGGAATCTCCCATTTATATGAGTCCTGAGGCTC 610  
Db 563 CAAAACCAAGGCTATGACACAAAGTGGACATCTGGTCACTGGCATCATTTTGTATG 670  
Qy 611 TGAAACACCAAGGCTATGACACAAAGTGGACATCTGGTCACTGGCATCATTTTGTATG 670  
Db 623 TACATGAAATGGATGATCACTTCAAACTGACATCTGGTCTTGGCTGTCTACTATATG 682  
Qy 671 AGATGCTGCTGCAATGATCATGCTTGGCTGGCTCCAA 707  
Db 683 AGATGGCTGCAATGATCAAAAGTCTTCTATGGTGACAA 719

RESULT 11  
US-09-221-527-3  
; Sequence 3, Application US/09221527  
; Patent No. 6146832  
; GENERAL INFORMATION:  
; APPLICANT: Acton, Susan  
; TITLE OF INVENTION: NOVEL CSAPK-1 NUCLEIC ACID MOLECULES AND USES THEREFOR  
; FILE REFERENCE: MNI-050  
; CURRENT APPLICATION NUMBER: US/09/221,527  
; EARLIER FILING DATE: 1998-12-28  
; EARLIER FILING DATE:  
; NUMBER OF SEQ ID NOS: 15  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 3  
; LENGTH: 906  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: (1)..(906)  
US-09-221-527-3

Query Match 5.9%; Score 85; DB 3; Length 906;  
Best Local Similarity 49.1%; Pred. No. 2.7e-15;  
Matches 254; Conservative 0; Mismatches 260; Indels 3; Gaps 1;  
Qy 194 TTGGAGAACTAAATCCAAATGAACTGTACAGGCCAATTTGGAGCCCACTCTCTCCA 253  
Db 203 TTGATTTAATGGATGCCAAGCAGCTGCTGATTCATCAAGAAATAGATCTTCTTAAGC 262  
Qy 254 AGCTGGACCAACCCAGCCATTTGTCAGTTCATGCAAGTTTGGAGCAAGATAATTTCT 313  
Db 263 AACTCAACCATCCAAATGTAATAAATATTTATGATCAATCAAGATATGAATGA 322  
Qy 314 GATTTACACGAGTACTGTGAGGCGGAGATCTGACGATATAAATTTTGGAGCAAGATAATTAAC 373  
Db 323 AACTCAACCATCCAAATGTAATAAATATTTATGATCAATCAAGATATGAATGA 382  
Qy 374 AAGCTGGAATAATCTTTCCAGAAATCAATAATAGAAATGGTTTATCCAGCTGCTGCTGG 433  
Db 383 AGCAAAAGAGGCTAAATCTCTGAAAGAACTGTTTGGAAATATTTTGTTCAGCTTTCAGTG 442  
Qy 434 GAGTTGACTACATGCGATGAGAGAGAGATCTTCATCGAGACTTAAAGTCAAGAAATGTAT 493  
Db 443 CATTTGAAACACATGCAATCTCGAAGAGTCAATGATAGATATAAAACCAAGTAAATGTGT 502  
Qy 494 TTCTGAAA---AATAATCTCTTTAAATTTGGAGATTTTGGAGTTTCTCGACTTCTAATGG 550  
Db 503 TCATTACAGCCACTGGGGTGTAAACTTGGAGATCTTGGCGCTTGGCGGTTTTTTCAGCT 562  
Qy 551 GATCCTGTGACCTGGCCACAACTTTAACTGGAATCTCCCATTTATATGAGTCTGAGGCTC 610  
Db 563 CAAAACCAAGCTGCAATCTTTAGTTGGTACGCTTATTACATGTCTTCCAGAGAA 622

QY 611 TGAACACCAAGCTATGACACAAAGTCGACATCTGGTCACTGGCATGCAATTTGTATG 670  
Db 623 TACATCAAAATGATACAACTTCAATCTGACATCTGGTCTCTTGGCTGTCTACTATATG 682  
QY 671 AGATGCTGCATGAATCATGATTCGCTGGCTCAA 707  
Db 683 AGATGCTGCATTAACAAAGTCCTTCTATGGTGACAA 719

## RESULT 12

US-09-221-236-3

; Sequence 3, Application US/09221236

; Patent No. 6146841

; GENERAL INFORMATION:

; APPLICANT: Acton, Susan

; TITLE OF INVENTION: NOVEL CSAPK-1 NUCLEIC ACID MOLECULES AND USES THEREFOR

; FILE REFERENCE: MNI-050

; CURRENT APPLICATION NUMBER: US/09/221,236

; EARLIER FILING DATE: 1998-12-28

; EARLIER APPLICATION NUMBER: 09/163,115

; EARLIER FILING DATE: 1998-09-29

; NUMBER OF SEQ ID NOS: 15

; SOFTWARE: PatentIn Ver. 2.0

; SEQ ID NO 3

; LENGTH: 906

; TYPE: DNA

; ORGANISM: Homo sapiens

; FEATURE:

; NAME/KEY: CDS

; LOCATION: (1)..(906)

; US-09-221-236-3

## Query Match

Best Local Similarity 5.9%; Score 85; DB 3; Length 906;

Matches 254; Conservative 0; Mismatches 260; Indels 3; Gaps 1;

QY 194 TTGGAGAACTAAATCCAAATGAACCTGTACAGCCCAATTTGGAAGCCCACTCTCTCCA 253  
Db 203 TTGATTAAATGATGCAAGACGCTGCTGATGTCATCAAGAAATAGATCTCTTAAGC 262

QY 254 AGCTGGACCAACCCAGCCATTTGCAAGTTCATGCAAGTTCCTGGAGCAAGATAATTTCT 313  
Db 263 AACTCAACCATCCAAATGTAATAAATAATTATGTCATCAATCGAATAATGAATAA 322

QY 314 GCATTTATCAGGAGTACTGTGAGGCGCGAGATCTGGAACGATAAAATTCAGGAATATAA 373  
Db 323 ACATAGTTTGGAACTAGCAGATCTCGGACCTATCCAGAATGATCAAGCATTTTAAGA 382

QY 374 AAGCTGGAAAAATCTTTCCAGAAAAATCAAAATAATAGATGGTTTATCCAGCTGCTGCTG 433  
Db 383 AGCAAAAGAGGCTAATTCCTGAAAGAACTGTTGGAAGTATTTTTCAGCTTTCAGTG 442

QY 434 GAGTTGACTACATGATGAGAGGAGATATCTTCATCGAGACTTAAAGTCAAAAGATGTAT 493  
Db 443 CATTTGGAACACATGCAATCTCGAAGAGTCATGCAATAGATATAAAACCAGCTAATGT 502

QY 494 TTCTGAAA---AATAATCTCTTAAATTTGAGATTTTGGAGTTTCTCGACTTCTAATGG 550  
Db 503 TCATTACAGCCACTGGGGTGGTAAACTTTGAGATCTTGGCTTGGCCCGTTTTCAGCT 562

QY 551 GATCTGTGACCTGGCCACAACTTTAACTGGAACCTCCCATTAATATGAGTCTGAGGCTC 610  
Db 563 CAAAACCAACGCTGCACATCTTTTAGTTGGTAGCGCTTATTACATGCTCCAGAGAGAA 622

QY 611 TGAACACCAAGCTATGACACAAAGTCGCAATCTGGTCACTGGCATGCAATTTGTATG 670  
Db 623 TACATGAAAATGGATACAACTTCAAAATCTGACATCTGGTCTCTTGGCTGTCTACTATATG 682

QY 671 AGATGCTGCATGAATCATGCAATTCGCTGGCTCCAA 707

Db 683 AGATGCTGCATTAACAAAGTCCTTCTATGGTGACAA 719

## RESULT 13

US-09-221-416-3

; Sequence 3, Application US/09221416

; Patent No. 6153417

; GENERAL INFORMATION:

; APPLICANT: Acton, Susan

; TITLE OF INVENTION: NOVEL CSAPK-1 NUCLEIC ACID MOLECULES AND USES THEREFOR

; FILE REFERENCE: MNI-050

; CURRENT APPLICATION NUMBER: US/09/221,416

; EARLIER FILING DATE: 1998-12-28

; EARLIER APPLICATION NUMBER: 09/163,115

; EARLIER FILING DATE: 1998-09-29

; NUMBER OF SEQ ID NOS: 15

; SOFTWARE: PatentIn Ver. 2.0

; SEQ ID NO 3

; LENGTH: 906

; TYPE: DNA

; ORGANISM: Homo sapiens

; FEATURE:

; NAME/KEY: CDS

; LOCATION: (1)..(906)

; US-09-221-416-3

## Query Match

Best Local Similarity 5.9%; Score 85; DB 3; Length 906;

Matches 254; Conservative 0; Mismatches 260; Indels 3; Gaps 1;

QY 194 TTGGAGAACTAAATCCAAATGAACCTGTACAGCCCAATTTGGAAGCCCACTCTCTCCA 253  
Db 203 TTGATTAAATGATGCAAGACGCTGCTGATGTCATCAAAAGAAATAGATCTCTTAAGC 262

QY 254 AGCTGGACCAACCCAGCCATTTGCAAGTTCATGCAAGTTCCTGGAGCAAGATAATTTCT 313  
Db 263 AACTCAACCATCCAAATGTAATAAATAATTATGTCATCAATCGAATAATGAATAA 322

QY 314 GCATTTATCAGGAGTACTGTGAGGCGCGAGATCTGGAACGATAAAATTCAGGAATATAA 373  
Db 323 ACATAGTTTGGAACTAGCAGATCTCGGACCTATCCAGAATGATCAAGCATTTTAAGA 382

QY 374 AAGCTGGAAAAATCTTTCCAGAAAAATCAAAATAATAGATGGTTTATCCAGCTGCTGCTG 433  
Db 383 AGCAAAAGAGGCTAATTCCTGAAAGAACTGTTGGAAGTATTTTTCAGCTTTCAGTG 442

QY 434 GAGTTGACTACATGATGAGAGGAGATATCTTCATCGAGACTTAAAGTCAAAAGATGTAT 493  
Db 443 CATTTGGAACACATGCAATCTCGAAGAGTCATGCAATAGATATAAAACCAGCTAATGT 502

QY 494 TTCTGAAA---AATAATCTCTTAAATTTGAGATTTTGGAGTTTCTCGACTTCTAATGG 550  
Db 503 TCATTACAGCCACTGGGGTGGTAAACTTTGAGATCTTGGCTTGGCCCGTTTTCAGCT 562

QY 551 GATCTGTGACCTGGCCACAACTTTAACTGGAACCTCCCATTAATATGAGTCTCGAGGCTC 610  
Db 563 CAAAACCAACGCTGCACATCTTTTAGTTGGTAGCGCTTATTACATGCTCCAGAGAGAA 622

QY 611 TGAACACCAAGGCTATGACACAAAGTCGCAATCTGGTCACTGGCATGCAATTTGTATG 670  
Db 623 TACATGAAAATGGATACAACTTCAAAATCTGACATCTGGTCTCTTGGCTGTCTACTATATG 682

QY 671 AGATGCTGCATGAATCATGCAATTCGCTGGCTCCAA 707

Db 683 AGATGCTGCATTAACAAAGTCCTTCTATGGTGACAA 719

## RESULT 14

US-09-221-245-3

; Sequence 3, Application US/09221245

; Patent No. 6180358

; GENERAL INFORMATION:

; APPLICANT: Acton, Susan

; TITLE OF INVENTION: NOVEL CSAPK-1 NUCLEIC ACID MOLECULES AND USES THEREFOR

; FILE REFERENCE: MNI-050

; CURRENT APPLICATION NUMBER: US/09/221,245

; CURRENT FILING DATE: 1998-12-28  
; EARLIER APPLICATION NUMBER: US 09/163,115  
; EARLIER FILING DATE: 1998-09-29  
; NUMBER OF SEQ ID NOS: 15  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 3  
; LENGTH: 906  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: (1)..(906)  
US-09-221-245-3

Query Match 5.9%; Score 85; DB 3; Length 906;  
Best Local Similarity 49.1%; Pred. No. 2.7e-15;  
Matches 254; Conservative 0; Mismatches 260; Indels 3; Gaps 1;  
QY 194 TTGGAGAACTAAATCCAAATGAAACTGTACAGGCCAAATTTGGAAGCCCAACTCCTCTCCA 253  
DB 203 TTGATTTAATGGATGCCAAGCAGCTGCTGATTGCATCAAGAAATAGATCTTCTTAAGC 262  
QY 254 AGCTGGACCAACCCAGCCATTGTCAAGTTCATGCAAGTTTGTGGAGCAAGATAATTTCT 313  
DB 263 AACTCAACCAATCCAAATGTAATAAATATATGATCATCAATGGAAGATAATGAACATA 322  
QY 314 GCATTATCACGGAGTACTGTGAGGGCGGAGATCTGGACGATATAAATTCAGGAATATAAAC 373  
DB 323 ACATAGTTTGGAACTAGCAGATGCTGGCGACCTATCCAGAAATGATCAAGCATTTTAAGA 382  
QY 374 AAGCTGGAAAAATCTTCCAGAAAAATCAAAATAATAGAAATGGTTTATCCAGCTGCTGCTG 433  
DB 383 AGCAAAAGAGGCTAATTCCTGAAAGAACTGTTTGGAAAGTATTTTGTTCAGCTTTCAGTG 442  
QY 434 GAGTTGACTACATGCTAGAGGAGGATCTTCATCGAGACTTAAAGTCAAAAGATGTAT 493  
DB 443 CATTTGAAACACATGCAATCTCGAAGAGTCATGCAATAGATATAAACCAGCTAATGTGT 502  
QY 494 TTCGAAAA---AATAATCTCTTAAATTTGGAAGTATTTTGGAAAGTATTTTGTTCAGCTTTCAGTG 550  
DB 503 TCATTACAGCCACTGGGGTGTAAACTTGGAGATCTTGGGCTTGGCCGGTTTTTCAGCT 562  
QY 551 GATCCTGTGACCTGGCCACAACTTAACTGGAAGTATTTGGAAGTATTTTGTTCAGCTTTCAGTG 610  
DB 563 CAAAACCAACAGCTGCACATCTTTAGTTGGTACGCTTATGATGCTTCCAGAGAGAA 622  
QY 611 TGAACACCAACAGCTATGACACAAAGTCGAGACATCTGGTCACTGGCATGCAATTTGTATG 670  
DB 623 TACATGAAATGGATACAACTTCAAAATCTGACATCTGGTCTCTTGGCTGTCTACTATATG 682  
QY 671 AGATGTGCTGCATGATCATGCAATTCGCTGGCTCCAA 707  
DB 683 AGATGGCTGCATTACAAAGTCTTCTATGGTGACAA 719

RESULT 15  
US-09-163-115-3  
; Sequence 3, Application US/09163115A  
; Patent No. 6183962  
; GENERAL INFORMATION:  
; APPLICANT: Acton, Susan  
; TITLE OF INVENTION: NOVEL CSAPK-1 NUCLEIC ACID MOLECULES AND USES THEREFOR  
; FILE REFERENCE: MNI-050  
; CURRENT APPLICATION NUMBER: US/09/163,115A  
; CURRENT FILING DATE: 1998-09-29  
; NUMBER OF SEQ ID NOS: 15  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 3  
; LENGTH: 906  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: CDS

; LOCATION: (1)..(906)  
US-09-163-115-3

Query Match 5.9%; Score 85; DB 3; Length 906;  
Best Local Similarity 49.1%; Pred. No. 2.7e-15;  
Matches 254; Conservative 0; Mismatches 260; Indels 3; Gaps 1;

QY 194 TTGGAGAACTAAATCCAAATGAAACTGTACAGGCCAAATTTGGAAGCCCAACTCCTCTCCA 253  
DB 203 TTGATTTAATGGATGCCAAGCAGCTGCTGATTGCATCAAGAAATAGATCTTCTTAAGC 262  
QY 254 AGCTGGACCAACCCAGCCATTGTCAAGTTCATGCAAGTTTGTGGAGCAAGATAATTTCT 313  
DB 263 AACTCAACCAATCCAAATGTAATAAATATATGATCATCAATGGAAGATAATGAACATA 322  
QY 314 GCATTATCACGGAGTACTGTGAGGGCGGAGATCTGGACGATATAAATTCAGGAATATAAAC 373  
DB 323 ACATAGTTTGGAACTAGCAGATGCTGGCGACCTATCCAGAAATGATCAAGCATTTTAAGA 382  
QY 374 AAGCTGGAAAAATCTTCCAGAAAAATCAAAATAATAGAAATGGTTTATCCAGCTGCTGCTG 433  
DB 383 AGCAAAAGAGGCTAATTCCTGAAAGAACTGTTTGGAAAGTATTTTGTTCAGCTTTCAGTG 442  
QY 434 GAGTTGACTACATGCTAGAGGAGGATCTTCATCGAGACTTAAAGTCAAAAGATGTAT 493  
DB 443 CATTTGAAACACATGCAATCTCGAAGAGTCATGCAATAGATATAAACCAGCTAATGTGT 502  
QY 494 TTCGAAAA---AATAATCTCTTAAATTTGGAAGTATTTTGGAAAGTATTTTGTTCAGCTTTCAGTG 550  
DB 503 TCATTACAGCCACTGGGGTGTAAACTTGGAGATCTTGGGCTTGGCCGGTTTTTCAGCT 562  
QY 551 GATCCTGTGACCTGGCCACAACTTAACTGGAAGTATTTTGGAAAGTATTTTGTTCAGCTTTCAGTG 610  
DB 563 CAAAACCAACAGCTGCACATCTTTAGTTGGTACGCTTATGATGCTTCCAGAGAGAA 622  
QY 611 TGAACACCAACAGGCTATGACACAAAGTCGAGACATCTGGTCACTGGCATGCAATTTGTATG 670  
DB 623 TACATGAAATGGATACAACTTCAAAATCTGACATCTGGTCTCTTGGCTGTCTACTATATG 682  
QY 671 AGATGTGCTGCATGATCATGCAATTCGCTGGCTCCAA 707  
DB 683 AGATGGCTGCATTACAAAGTCTTCTATGGTGACAA 719

Search completed: October 3, 2004, 00:58:23  
Job time : 122 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: October 2, 2004, 06:03:38 ; Search time 612 Seconds

(without alignments)  
10058.234 Million cell updates/sec

Title: US-10-803-278-3

Perfect score: 1449

Sequence: 1 atgtggaatcccaagagc.....actggaaaatgaaaataa 1449

Scoring table: IDENTITY\_NUC

Gapop 10.0 , Gapext 1.0

Searched: 3373863 seqs, 2124099041 residues

Total number of hits satisfying chosen parameters: 6747726

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : N\_Geneseq\_29Jan04.\*  
1: Geneseqn1980s.\*  
2: Geneseqn1990s.\*  
3: Geneseqn2000s.\*  
4: Geneseqn2001as.\*  
5: Geneseqn2001bs.\*  
6: Geneseqn2002s.\*  
7: Geneseqn2003as.\*  
8: Geneseqn2003bs.\*  
9: Geneseqn2003cs.\*  
10: Geneseqn2004s.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1399.8	96.6	2080	9	ADCS1183 Human cel
2	1399	96.5	1938	4	AAS06746 Polynucle
3	1399	96.5	2552	6	AAD30552 Human kin
4	1399	96.5	2939	9	ADCS1185 Human cel
5	1397.4	96.4	2869	6	ABK89295 Human cDN
6	1385.4	95.6	2241	6	ABZ11168 Human pol
7	1179	81.4	2888	9	ADCS9943 Human KPP
8	1063	73.4	2631	10	ADCE28343 Human KPP
9	759	52.4	2483	9	ADCS99142 Human KPP
10	669.6	46.2	2538	6	ABL34984 Rat cDNA
11	535	36.9	1071	9	ADCS1181 Human cel
12	415.8	28.7	451	4	ABK43470 DNA encod
13	395.4	27.3	430	4	AAS27206 cDNA enco
14	395.4	27.3	430	4	AAK88454 Human dig
15	395.4	27.3	430	4	ABK43786 DNA encod
16	395.4	27.3	430	9	ADB93384 Human cDN
17	336	23.2	336	9	ADCS1179 Human cel
18	199	13.7	1438	4	AAI02041 Human rep
19	173.6	12.0	287	6	ABN25194 Human ORF
20	172.2	11.9	326014	6	ABK89296 Human gen
21	156	10.8	2022	4	AAI04664 Human rep
22	156	10.8	2022	4	ABJ97571 Human tes
23	156	10.8	2961	4	AAI04665 Human rep

C

## ALIGNMENTS

### RESULT 1

ADCS1183  
ID ADCS1183 standard; DNA; 2080 BP.

XX AC ADCS1183;

XX DT 18-DEC-2003 (first entry)

XX DE Human cell-cycle related protein coding sequence, SEQ ID 9.

XX KW Human; cytostatic; cell-cycle related protein; nuclear export;  
XX KW nuclear-cytoplasm transport; cytotoxic; cell-cycle control;  
XX KW immunological disease; neurological disease; cancer; gene; ds.

XX OS Homo sapiens.

XX FH Key Location/Qualifiers  
XX CDS 294..1706

XX FT /\*tag= a  
XX FT /product= "Cell-cycle related protein"

XX PN JP2003144168-A.

XX PD 20-MAY-2003.

XX PF 14-NOV-2001; 2001JP-00349158.

XX PR 14-NOV-2001; 2001JP-00349158.

XX PA (KOKU-) KOKURITSU YOBO EISEI KENKYUSHO.

XX PA (KAGA-) KAGAKO GIJUTSU SHINKO JIGYODAN.

XX DR WPI; 2003-818166/77.

XX DR P-PSDB; ADCS1184.

XX PT Novel DNA or RNA coding a cell-cycle related protein which has nuclear  
XX PT export function, useful for screening substance that prevent or treat  
XX PT cell cycle abnormality diseases e.g. immunological disease.

XX PS Disclosure; SEQ ID NO 9; 41pp; Japanese.

XX CC The present invention relates to novel cell-cycle related protein such as  
XX CC NIMA (Never-In Mitosis, gene A)-related protein kinase of Nek 9 and  
XX CC coding sequences such as a cell-cycle related protein (ADCS1176) having  
XX CC nuclear export function; cell-cycle related protein (ADCS1178) having  
XX CC nuclear-cytoplasm transport function; cell-cycle related protein

Ab197572 Human tes  
Abx72260 Human NOV  
Abx72259 Human NOV  
Abx72261 Human NOV  
Abz68734 Nucleotid  
Abz68773 Nucleotid  
Abz68735 Nucleotid  
Abz77150 Human pro  
Adc99119 Human KPP  
Aaz46147 cDNA sequ  
Abz59800 Human ser  
Aad38850 Human kin  
Aak51477 Human pol  
Abz59718 NEK-like  
Abz68774 Nucleotid  
Abns9637 Novel hum  
Aak52461 Human pol  
Aaa09328 Human can  
Aas90554 DNA encod  
Abv75946 Ser/Thr/T  
Aas11558 Human cDN  
Aal51590 Human ser

24 156 10.8 2961 4 ABL97572  
25 149.4 10.3 1781 7 ABX72260  
26 149.4 10.3 2257 7 ABX72259  
27 148.2 10.2 1588 7 ABX72261  
28 147.8 10.2 1510 7 ABZ68734  
29 147.8 10.2 1578 7 ABZ68773  
30 147.8 10.2 1581 7 ABZ68735  
31 147.8 10.2 1638 7 ABZ77150  
32 147.8 10.2 2050 9 ADC99119  
33 147.8 10.2 2059 3 AAZ46147  
34 147.8 10.2 2102 6 ABZ69800  
35 147.8 10.2 2110 6 AAD38850  
36 147.8 10.2 2220 4 AAK51477  
37 146.4 10.1 2715 7 ABZ59718  
38 146.4 10.1 2715 7 ABZ68774  
39 146.2 10.1 2354 6 ABNS9637  
40 146.2 10.1 2372 4 AAK52461  
41 146.2 10.1 4253 3 AAA09328  
42 144 9.9 1403 5 AAS90554  
43 143.2 9.9 2343 6 ABV75946  
44 143.2 9.9 3645 5 AAS11558  
45 143.2 9.9 3729 7 AAL51590

C

(ADCS1180) having nuclear export function, cytotoxic function and transfer function in the nucleus; and/or cell-cycle related protein (ADCS1182) having transfer function and cytotoxic function in nucleus, nuclear export function and nuclear cytoplasm transport function. The coding sequences for these proteins are given in ADCS1177, ADCS1179 and ADCS1181. The sequences of the invention are useful for screening a substance which promotes or suppresses the transfer function or cytotoxic function in the nucleus, nuclear-cytoplasm transport function and/or a cell-cycle control function. The sequences are also useful for treating or diagnosing an immunological disease, neurological disease or cancer. The present sequence was used to illustrate the invention.

XX Sequence 2080 BP; 652 A; 439 C; 473 G; 516 T; 0 U; 0 Other;

Query Match 96.6%; Score 1399.8; DB 9; Length 2080;  
 Best Local Similarity 99.2%; Pred. No. 0;  
 Matches 1407; Conservative 0; Mismatches 12; Indels 0; Gaps 0;

QY 1 ATGCTGAATTCGAGGAGCTAAGTGTGTGAGTGGATCAACAGCCATTTCCACTTAT 60  
 DB 294 ATGCTGAATTCGAGGAGCTAAGTGTGTGAGTGGATCAACAGCCATTTCCACTTAT 353  
 QY 61 CCAAGACCTTGAATTCGAGGAGCTAAGTGTGTGAGTGGATCAACAGCCATTTCCACTTAT 120  
 DB 354 CCAAGACCTTGAATTCGAGGAGCTAAGTGTGTGAGTGGATCAACAGCCATTTCCACTTAT 413  
 QY 121 GGAATCTGTATCTGTTTCAAGAGAAAGCAACGAGGAGAGGAATTAAGTACTT 180  
 DB 414 GGAATCTGTATCTGTTTCAAGAGAAAGCAACGAGGAGAGGAATTAAGTACTT 473  
 QY 181 ARGGAATATCTGTTGGAGAACTTAATCCAAATGAACGTACAGCCCAATTTGCAAGCC 240  
 DB 474 ARGGAATATCTGTTGGAGAACTTAATCCAAATGAACGTACAGCCCAATTTGCAAGCC 533  
 QY 241 CAACTCTCTCCAAAGCTGGACCAACCCAGCCATTTGCAAGTTCCTCAAGTTCCTGGAG 300  
 DB 534 CAACTCTCTCCAAAGCTGGACCAACCCAGCCATTTGCAAGTTCCTCAAGTTCCTGGAG 593  
 QY 301 CAAGATAATTTGCAATATCAGGAGTACTGTGAGGCGGAGATCTGGACGATAAAT 360  
 DB 594 CAAGATAATTTGCAATATCAGGAGTACTGTGAGGCGGAGATCTGGACGATAAAT 653  
 QY 361 CAGGAATATAACAGCTGGAATAATCTTTCCAGAAATCAATTAATAGAATGGTTTATC 420  
 DB 654 CAGGAATATAACAGCTGGAATAATCTTTCCAGAAATCAATTAATAGAATGGTTTATC 713  
 QY 421 CAGCTCTCTGGAGTTCATCATGATGATGAGAGGATACCTTCATCGAGACTTAAAG 480  
 DB 714 CAGCTCTCTGGAGTTCATCATGATGATGAGAGGATACCTTCATCGAGACTTAAAG 773  
 QY 481 TCAAGAAGTATTTCTGGAATAATCTCTTAAATTTGAGATTTGGAGTTTCTCGA 540  
 DB 774 TCAAGAAGTATTTCTGGAATAATCTCTTAAATTTGAGATTTGGAGTTTCTCGA 833  
 QY 541 CTTCTAATGGATCTCTGACCTGGCCCAACATTTAATGGAATCTCCCATPATATGAT 600  
 DB 834 CTTCTAATGGATCTCTGACCTGGCCCAACATTTAATGGAATCTCCCATPATATGAT 893  
 QY 601 CTTGAGGCTCTGAACACCAAGCTATGACACAAAGTCGACATCTGGCTGCGATGC 660  
 DB 894 CTTGAGGCTCTGAACACCAAGCTATGACACAAAGTCGACATCTGGCTGCGATGC 953  
 QY 661 ATTGTTGATGAGATGTGTCATGATCATGCAATTCGCTGGCTCCCAATTTCTATPCCAT 720  
 DB 954 ATTGTTGATGAGATGTGTCATGATCATGCAATTCGCTGGCTCCCAATTTCTATPCCAT 1013  
 QY 721 GTTTTAAATTTGTTGAAGTGACACACCTTCTCTCCCTGAGAGATATCCAAAGACTA 780  
 DB 1014 GTTTTAAATTTGTTGAAGTGACACACCTTCTCTCCCTGAGAGATATCCAAAGACTA 1073  
 QY 781 AATGCCATCATGGAAGCATGTTGAACAAAGATCTCTTCAATTAAGACCATCTGCTATCGAA 840

Db 1074 AATGCCATCATGGAAGCATGTTGAACAAAGATCTCTTCAATTAAGACCATCTGCTATCGAA 1133  
 QY 841 ATTTTAAATTTCCCTTACCTTACCTTACCTTACCTTACCTTACCTTACCTTACCTTACCTT 900  
 Db 1134 ATTTTAAATTTCCCTTACCTTACCTTACCTTACCTTACCTTACCTTACCTTACCTT 1193  
 QY 901 ATGACTCTGGAAGCAAAAAATTTGGATTTGATGTTGATGTTGATGTTGATGTTGATGTTGATGTT 960  
 Db 1194 ATGACTCTGGAAGCAAAAAATTTGGATTTGATGTTGATGTTGATGTTGATGTTGATGTTGATGTT 1253  
 QY 961 ATGCAAAAAAGGATCCACCTGCGAGCTCTGAGGAGCTGTCAGAGTACAGAAATTAAGC 1020  
 Db 1254 ATGCAAAAAAGGATCCACCTGCGAGCTCTGAGGAGCTGTCAGAGTACAGAAATTAAGC 1313  
 QY 1021 CCAAGAGAAAGGATGCGGCTGAGGAAAGCTCCAGCGGCTGATCAGAAAGCCAGGAAAGCTG 1080  
 Db 1314 CCAAGAGAAAGGATGCGGCTGAGGAAAGCTCCAGCGGCTGATCAGAAAGCCAGGAAAGCTG 1373  
 QY 1081 AAAAAAGATTTGGAAGAAATATGAGAAATAGCAAAACGATGCAAGATTTGAGATCT 1140  
 Db 1374 AAAAAAGATTTGGAAGAAATATGAGAAATAGCAAAACGATGCAAGATTTGAGATCT 1433  
 QY 1141 CGGAACCTTTGAGAGCTGAGTGTGATGTTGATGTTGATGTTGATGTTGATGTTGATGTTGATGTT 1200  
 Db 1434 CGGAACCTTTGAGAGCTGAGTGTGATGTTGATGTTGATGTTGATGTTGATGTTGATGTTGATGTT 1493  
 QY 1201 GAAGAAAGGAGGAGCACTGAGGAGAGCTTTCTGTTTCTGTTTCTGTTTCTGTTTCTGTTTCTGTTT 1260  
 Db 1494 GAAGAAAGGAGGAGCACTGAGGAGAGCTTTCTGTTTCTGTTTCTGTTTCTGTTTCTGTTTCTGTTT 1553  
 QY 1261 GAGAGGTGGCAAGCGAGGAGAGGATCTGATGAAACCAACTTTAGAGAACCTGCTGAG 1320  
 Db 1554 GAGAGGTGGCAAGCGAGGAGAGGATCTGATGAAACCAACTTTAGAGAACCTGCTGAG 1613  
 QY 1321 TCTCAGCTTATCTTCCATGAGACCTCCAGAACCTTGAATCAATTTGAGGATGCCACA 1380  
 Db 1614 TCTCAGCTTATCTTCCATGAGACCTCCAGAACCTTGAATCAATTTGAGGATGCCACA 1673  
 QY 1381 TCTGACCTTCGATACCATGAGACTGTAATCTAATTTCA 1419  
 Db 1674 TCTGACCTTCGATACCATGAGACTGTAATCTAATTTCA 1712

RESULT 2  
 AAS06746  
 ID AAS06746 standard; cdna; 1938 BP.  
 XX  
 AC AAS06746;  
 XX  
 DT 12-SEP-2001 (first entry)  
 XX  
 DE Polynucleotide sequence encoding human protein kinase #46.  
 XX  
 KW Human; protein kinase; PK; STK; cancer; cardiovascular disease;  
 KW metabolic disorder; immune related disease; neurological disorder;  
 KW neurodegenerative disorder; inflammatory disorder; infectious disease;  
 KW reproductive disorder; gene therapy; ss.  
 XX  
 OS Homo sapiens.  
 XX  
 FN W0200138503-A2.  
 XX  
 PD 31-MAY-2001.  
 XX  
 PF 22-NOV-2000; 2000WO-US032085.  
 XX  
 PR 24-NOV-1999; 99US-0167482P.  
 XX  
 PA (SUGEN-) SUGEN INC.  
 XX  
 PI Plowman CD, Whyte D, Manning G, Sudarsanam S, Martinez R;  
 PI Flanagan P, Clary D;  
 XX



DR WPI; 2001-343950/36.  
DR P-PSDB; AAU03546.  
XX  
PT Nucleic acids encoding human kinase polypeptides, useful for preventing  
PT diagnosing and/or treating e.g. cancer, immune, cardiovascular and  
PT neuronal-associated diseases, and microbial infections.  
XX  
PS  
PS Example 1; Fig 1; 433pp; English.  
XX  
CC AAS06701-AAS06757 encode for novel human protein kinases #1-57. The novel  
CC protein kinases have been identified as members of the tyrosine or  
CC serine/threonine kinase (PTK and STK) families. The polynucleotides  
CC encoding protein kinases and the polypeptides may be used in the  
CC prevention, diagnosis and treatment of diseases associated with  
CC inappropriate kinase expression. For example, they may be used to treat  
CC cancers (especially cancers of haematopoietic origin), cardiovascular  
CC disease (e.g. atherosclerosis), metabolic disorders (e.g. diabetes),  
CC immune related diseases (e.g. rheumatoid arthritis), neurological  
CC disorders (e.g. schizophrenia), neurodegenerative disorders (e.g.  
CC Parkinson's disease), inflammatory disorders (e.g. asthma), infectious  
CC disease (e.g. HIV) and reproductive disorders (e.g. infertility).  
CC Additionally, polynucleotides encoding protein kinases may be used for  
CC gene therapy, and as DNA probes in diagnostic assays. The protein kinase  
CC polypeptides may be used as antigens in the production of antibodies  
CC against the protein kinases and in assays to identify modulators of  
CC protein kinase expression and activity  
XX  
SQ Sequence 1938 BP; 646 A; 383 C; 461 G; 448 T; 0 U; 0 Other;  
Query Match 96.5%; Score 1399; DB 4; Length 1938;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 1399; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 ATGCTGAAATCCAGAGCAGCTAAGTGTGTGAGTGGATCAACGCCATTTCCTAT 60  
DB 1 ATGCTGAAATCCAGAGCAGCTAAGTGTGTGAGTGGATCAACGCCATTTCCTAT 60  
QY 61 CCAAGACCTGATTCAGAGATACGTGCTTCAACAAAATCTGGCAGTGGAGTTT 120  
DB 61 CCAAGACCTGATTCAGAGATACGTGCTTCAACAAAATCTGGCAGTGGAGTTT 120  
QY 121 GGAATCTGTATCTGTTTCAGACAAAGAACCCAGAGGAGAGAAATTAAGGTACT 180  
DB 121 GGAATCTGTATCTGTTTCAGACAAAGAACCCAGAGGAGAGAAATTAAGGTACT 180  
QY 181 AAGGAATATCTGTTGGAGAACTAAATCCAAATGAACTGTACAGCCAAATTTGAGCC 240  
DB 181 AAGGAATATCTGTTGGAGAACTAAATCCAAATGAACTGTACAGCCAAATTTGAGCC 240  
QY 241 CAACTCTCTCCAGCTGGACCCAGCCATGTCAAGTTCATGCAAGTTTGTGGAG 300  
DB 241 CAACTCTCTCCAGCTGGACCCAGCCATGTCAAGTTCATGCAAGTTTGTGGAG 300  
QY 301 CAAAGATAATTTCTGCATTATCAGGAGTACTGTGAGGGCCGAGATCTGACAGNATAAAT 360  
DB 301 CAAAGATAATTTCTGCATTATCAGGAGTACTGTGAGGGCCGAGATCTGACAGNATAAAT 360  
QY 361 CAGGAATATAACAGCTGAAAAATCTTCCAGAAAAATCAATAATAGATGGTTTATC 420  
DB 361 CAGGAATATAACAGCTGAAAAATCTTCCAGAAAAATCAATAATAGATGGTTTATC 420  
QY 421 CAGCTGCTGCTGGAGTTGACTACATGATGATGAGAGGAGGATCTTCATCGAGACTTAAAG 480  
DB 421 CAGCTGCTGCTGGAGTTGACTACATGATGATGAGAGGAGGATCTTCATCGAGACTTAAAG 480  
QY 481 TCAAGATGATGTTCTGAAAAATATCTCTTAAATTTGGAGATTTGGAGTTTCTCGA 540  
DB 481 TCAAGATGATGTTCTGAAAAATATCTCTCTTAAATTTGGAGATTTGGAGTTTCTCGA 540  
QY 541 CTCTTAATGGATCTCTGACCTGGCCACAACTTTAACTGGAATCTCCCATATATAGT 600  
DB 541 CTCTTAATGGATCTCTGACCTGGCCACAACTTTAACTGGAATCTCCCATATATAGT 600

QY 601 CCTGAGGCTCTGAAACACCAAGGCTATGACACAAAAGTCGACATCTGGTCACTGGCATGC 660  
DB 601 CCTGAGGCTCTGAAACACCAAGGCTATGACACAAAAGTCGACATCTGGTCACTGGCATGC 660  
QY 661 ATTTTGTATGAGTGTCTGCATGAATCATGCAATTCGCTGCTCCCAATTTCTTATCCATT 720  
DB 661 ATTTTGTATGAGTGTCTGCATGAATCATGCAATTCGCTGCTCCCAATTTCTTATCCATT 720  
QY 721 GTTTTAAAAATTTGTAAGGTGACACACCTTCTCTCCCTGAGAGATATCAAAAGAACTA 780  
DB 721 GTTTTAAAAATTTGTAAGGTGACACACCTTCTCTCCCTGAGAGATATCAAAAGAACTA 780  
QY 781 AATGCCATCATGGAAGCATGTTGAAACAAGATCCCTTATTAGACCATCTGCTATCGAA 840  
DB 781 AATGCCATCATGGAAGCATGTTGAAACAAGATCCCTTATTAGACCATCTGCTATCGAA 840  
QY 841 ATTTTAAAAATCCCTTACCTTGTATGACAGCTACAGAACCTTAATGTGTAGATATTTCAGAA 900  
DB 841 ATTTTAAAAATCCCTTACCTTGTATGACAGCTACAGAACCTTAATGTGTAGATATTTCAGAA 900  
QY 901 ATGACTCTGGAAGCAAAAATTTGGATGTGACAGAGGAGCTGCTCATATATTAATGTC 960  
DB 901 ATGACTCTGGAAGCAAAAATTTGGATGTGACAGAGGAGCTGCTCATATTAATTAATGTC 960  
QY 961 ATGCAAAAAGGATCCACCTGCAGACTCTGAGGGCCTGTGAGAAATACAGAAAATGACG 1020  
DB 961 ATGCAAAAAGGATCCACCTGCAGACTCTGAGGGCCTGTGAGAAATACAGAAAATGACG 1020  
QY 1021 CCAAGAGAAAGGATGCGGCTGAGAAAGCTCCAGCGGCTGATGAGAAAGCCAGGAAGCTG 1080  
DB 1021 CCAAGAGAAAGGATGCGGCTGAGAAAGCTCCAGCGGCTGATGAGAAAGCCAGGAAGCTG 1080  
QY 1081 AAAAAGATTTGGAAGAAAAATATGAAGAAATAGCAAAAGATGCAAGAAATTCAGATCT 1140  
DB 1081 AAAAAGATTTGGAAGAAAAATATGAAGAAATAGCAAAAGATGCAAGAAATTCAGATCT 1140  
QY 1141 CGGAATCTTTCAGAGCTGAGTGTGATGTACTCCATGAAAAACACATTTAAAGGAATG 1200  
DB 1141 CGGAATCTTTCAGAGCTGAGTGTGATGTACTCCATGAAAAACACATTTAAAGGAATG 1200  
QY 1201 GAAAGAAAGGAGGAGCACTGAGGAGAGCTTTCTGTTTCAACCCAGGAGGAGATGAA 1260  
DB 1201 GAAAGAAAGGAGGAGCACTGAGGAGAGCTTTCTGTTTCAACCCAGGAGGAGATGAA 1260  
QY 1261 GAGAGGTGGCAAGCAGGAGAGGAAATCTGATCAACCACTTTAGAGAACTCTCCCTGAG 1320  
DB 1261 GAGAGGTGGCAAGCAGGAGAGGAAATCTGATCAACCACTTTAGAGAACTCTCCCTGAG 1320  
QY 1321 TCTCAGCCTATTCCTTCCATGGAACCTCCACGAATTCGAATTCAGAGGATGCCACA 1380  
DB 1321 TCTCAGCCTATTCCTTCCATGGAACCTCCACGAATTCGAATTCAGAGGATGCCACA 1380  
QY 1381 TCTGACCTTGGATACCATG 1399  
DB 1381 TCTGACCTTGGATACCATG 1399  
RESULT 3  
AAD30552  
ID AAD30552 standard; cDNA; 2552 BP.  
XX  
AC AAD30552;  
XX  
DT 21-MAY-2002 (first entry)  
XX  
DE Human kinase polypeptide (PKIN-5) cDNA.  
XX  
KW Human; kinase polypeptide; PKIN-5; gene therapy; Addison's disease;  
KW leukaemia; immune disorder; lymphoma; melanoma; developmental disorder;  
KW acquired immunodeficiency syndrome; AIDS; allergy; anaemia; hypertension;  
KW asthma; Crohn's disease; rheumatoid arthritis; bursitis; atherosclerosis;  
KW cirrhosis; hepatitis; psoriasis; Cushing's syndrome; cytostatic; cancer;  
KW cholestasis; cardiac; cardiovascular disorder; Niemann-Pick's disease;

lipid disorder; fatty liver; Gaucher's disease; myocardial infarction;  
drug screening; transgenic animal; antiinflammatory; hepatotropic;  
hypotensive; anti-HIV; enzyme; ss.

Homo sapiens.

Key Location/Qualifiers  
CDS 278..2215  
/\*tag= a  
/product= "Human PKIN-5"

W0200208399-A2.

31-JAN-2002.

20-JUL-2001; 2001WO-US023092.

21-JUL-2000; 2000US-0220038P.

28-JUL-2000; 2000US-0222112P.

04-AUG-2000; 2000US-0222831P.

11-AUG-2000; 2000US-0224729P.

(INCY-) INCYTE GENOMICS INC.

(THOR/) THORNTON M.

Thornton M, Yue H, Khan FA, Gururajan R, Hafalia AJA, Walia NK;  
Patterson C, Ramkumar J, Gandhi AR, Policky JL, Baughn MR;  
Tribouley CM, Bandman O, Nguyen DB, Lu Y, Burford N, Lal P, Ding L;  
Yao MG, Elliott VS, Recipon SA, Kearney L, Lu DAM, Greenwald SR;  
Tang YT, Xu Y, Walsh RT, Gietzen KU, Yang J, Hillman JL;

WPI; 2002-206083/26.

P-PSDB; AA619147.

New human kinase polypeptide, useful in diagnosis, prevention and  
treatment of cancer, immune disorder, growth and developmental disorder,  
cardiovascular disorder and lipid disorder.

Claim 5; Page 178-179; 196pp; English.

The present invention relates to an isolated human kinase polypeptide  
(PKIN) or a biologically active/immunogenic fragment of PKIN. PKIN is  
useful for diagnosing, treating and preventing cancer (e.g., leukemia,  
lymphoma, melanoma), an immune disorder (e.g., acquired immunodeficiency  
syndrome (AIDS), Addison's disease, allergy, anaemia, asthma, Crohn's  
disease, rheumatoid arthritis), a growth and developmental disorder (e.g.  
burstis, cirrhosis, hepatitis, psoriasis, Cushing's syndrome), a  
cardiovascular disorder (e.g., atherosclerosis, hypertension, myocardial  
infarction), and a lipid disorder (e.g., fatty liver, cholestasis,  
Gaucher's disease, Niemann-Pick's disease). PKIN is useful in a number of  
drug screening techniques and to analyse the proteome of a tissue or cell  
type. PKIN is useful for creating knockin humanised animals or transgenic  
animals to model human diseases, in somatic or germline gene therapy, to  
generate a transcript image of a tissue or cell type, for detecting  
differences in the chromosomal location due to translocation, inversion,  
etc., among normal, carrier or affected individuals, and as hybridisation  
probes for mapping naturally occurring genomic sequences. PKIN is useful  
in southern or northern analysis, dot blot or other membrane-based  
technologies, in PCR technologies, in dipstick, pin, microarray enzyme  
linked immunosorbent (ELISA)-like assays and in microarrays utilising  
fluids or tissues from patients to detect altered PKIN expression. The  
present sequence is human PKIN-5 cDNA

SQ Sequence 2552 BP; 813 A; 516 C; 608 G; 615 T; 0 U; 0 Other;

Query Match

Best Local Similarity 96.5%; Score 1399; DB 6; Length 2552;

Matches 1399; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGCTGAATTCAGAGGAGCTAAGTGTGTGAGTGGATCAACAGCCATTTCACCTTAT 60  
DB 278 ATGCTGAATTCAGAGGAGCTAAGTGTGTGAGTGGATCAACAGCCATTTCACCTTAT 337

QY 51 CCAGAGCCTTGATTGTCGAGGAGATAGTCTTCAACAAAAAAGCTTGGAGTGGAGTTT 120  
DB 338 CCAAGAGCCTTGATTGTCGAGGAGATAGTCTTCAACAAAAAAGCTTGGAGTGGAGTTT 397  
QY 121 GGAACCTGTCTATCTGTTTTCAGACAAAGAACCCAAAGAGGAGAGGAATTAAGAGTACTT 180  
DB 398 GGAACCTGTCTATCTGTTTTCAGACAAAGAACCCAAAGAGGAGGAATTAAGAGTACTT 457  
QY 181 AAGGAATATCTGTTGGAGAACTAAATCCAAATGAATGTACAGGCCAATTTGGAGGCC 240  
DB 458 AAGGAATATCTGTTGGAGAACTAAATCCAAATGAATGTACAGGCCAATTTGGAGGCC 517  
QY 241 CAATCTCTCTCAAGCTGGACCCAGCCATTTGTCAGATTCCTCAAGTTCCATGCAAGTTTGTGGAG 300  
DB 518 CAATCTCTCTCAAGCTGGACCCAGCCATTTGTCAGATTCCTCAAGTTCCATGCAAGTTTGTGGAG 577  
QY 301 CAAGATAATTTCTGCAATATCACCGAGTACTGTGAGGGCCGAGATCTGGACGATAAAAT 360  
DB 578 CAAGATAATTTCTGCAATATCACCGAGTACTGTGAGGGCCGAGATCTGGACGATAAAAT 637  
QY 361 CAGGAATATAAACAAGCTGGAAAAATCTTTCCAGAAAAATCAAAATAATAGAAATGTTTATC 420  
DB 638 CAGGAATATAAACAAGCTGGAAAAATCTTTCCAGAAAAATCAAAATAATAGAAATGTTTATC 697  
QY 421 CAGCTGCTGCTGGAGTGTGACTACATGATGAGAGGAGGATCTTCATCGAGACTTAAAG 480  
DB 698 CAGCTGCTGCTGGAGTGTGACTACATGATGAGAGGAGGATCTTCATCGAGACTTAAAG 757  
QY 481 TCRAAGATGTTTCTGAAAAATATCTCTTAATAATTTGGAGATTTTGGAGTTTCTCGA 540  
DB 758 TCRAAGATGTTTCTGAAAAATATCTCTTAATAATTTGGAGATTTTGGAGTTTCTCGA 817  
QY 541 CTTCTAATGGGATCCTCTGACCTGGCCACAACTTTAACTGGAACCTCCCACTTATATGAGT 600  
DB 818 CTTCTAATGGGATCCTCTGACCTGGCCACAACTTTAACTGGAACCTCCCACTTATATGAGT 877  
QY 601 CCTGAGCTCTGAACACCAAGGATGATGACAAAGTCGAGATCTGGTCACTGGGATGC 660  
DB 878 CCTGAGCTCTGAACACCAAGGATGATGACAAAGTCGAGATCTGGTCACTGGGATGC 937  
QY 661 ATTTTGTATGAGATGCTCTGCAATCATGATCATGCTGGCTCCCAATTTCTTATCCATT 720  
DB 938 ATTTTGTATGAGATGCTCTGCAATCATGATCATGCTGGCTCCCAATTTCTTATCCATT 997  
QY 721 GTTTTAAAAATTTGTTGAAGGTGACACACTTCTCTCCCTGAGAGATATCCAAAGAACTA 780  
DB 998 GTTTTAAAAATTTGTTGAAGGTGACACACTTCTCTCCCTGAGAGATATCCAAAGAACTA 1057  
QY 781 AATGCCATCATGAAAGCAATGTTGAAACAAAGATCTCTTCAATTAAGACCATCTGCTATCGAA 840  
DB 1058 AATGCCATCATGAAAGCAATGTTGAAACAAAGATCTCTTCAATTAAGACCATCTGCTATCGAA 1117  
QY 841 ATTTTAAAAATTCCTTACCTTGTATGAGAGCTACAGAACCTAATGTAGATATTCAGAA 900  
DB 1118 ATTTTAAAAATTCCTTACCTTGTATGAGAGCTACAGAACCTAATGTAGATATTCAGAA 1177  
QY 901 ATGACTCTGGAAGCAAAAAATTTGGATTGTGCAAGAGGCTGCTCATATAATATGTC 960  
DB 1178 ATGACTCTGGAAGCAAAAAATTTGGATTGTGCAAGAGGCTGCTCATATAATATGTC 1237  
QY 961 ATGCAAAAAAGGATCCACCTGACAGCTCTGAGGGCACTGTGAGAGGAGGAGTGAAGT 1020  
DB 1238 ATGCAAAAAAGGATCCACCTGACAGCTCTGAGGGCACTGTGAGAGGAGTGAAGT 1297  
QY 1021 CCAAGAGAAAGGATGGGCTGAGGAGAGCTCCAGGCGGCTGATGAGAAAGCCAGAGCTG 1080  
DB 1298 CCAAGAGAAAGGATGGGCTGAGGAGAGCTCCAGGCGGCTGATGAGAAAGCCAGAGCTG 1357  
QY 1081 AAAAAGATTTCTGGAAGAAAAATATGAAGAAAAATAGCAAAACGAATGCAAGATTTGAGATCT 1140  
DB 1358 AAAAAGATTTCTGGAAGAAAAATATGAAGAAAAATAGCAAAACGAATGCAAGATTTGAGATCT 1417  
QY 1141 CGGAACCTTTGAGGAGCTGAGTGTGATGTACTCCATGAAAAAACACATTTAAAGGAATG 1200



Db 1194 ATGACTCTGGAGACAAAATTTGGATTGTGAGAGGAGGCTGCTCATATTAATATGCC 1253  
 Qy 961 ATGCAAAAAGGATCCACCTGACAGACTCTGAGGCACTGTGAGAGTACAGAAAATGAGC 1020  
 Db 1254 ATGCAAAAAGGATCCACCTGACAGACTCTGAGGCACTGTGAGAGTACAGAAAATGAGC 1313  
 Qy 1021 CCAAGAGAAAGGATCGGCTGAGAACTCCAGGCGCTGATGAGAAAGCCAGAGCTG 1080  
 Db 1314 CCAAGAGAAAGGATCGGCTGAGAACTCCAGGCGCTGATGAGAAAGCCAGAGCTG 1373  
 Qy 1081 AAAAAGATTTGTGGAAGAAAATATGAAGAAAATAGCAAAATGCAAGAAATGAGATCT 1140  
 Db 1374 AAAAAGATTTGTGGAAGAAAATATGAAGAAAATAGCAAAATGCAAGAAATGAGATCT 1433  
 Qy 1141 CGGAACCTTCAGCAGCTGAGTGTGATGTACTCCATGAAAAACACATTTAAAGGAATG 1200  
 Db 1434 CGGAACCTTCAGCAGCTGAGTGTGATGTACTCCATGAAAAACACATTTAAAGGAATG 1493  
 Qy 1201 GAAGAAAAGGAGGACCACTGAGGGAAGACTTCTTGTTCACCCAGGACGAGGATGAA 1260  
 Db 1494 GAAGAAAAGGAGGACCACTGAGGGAAGACTTCTTGTTCACCCAGGACGAGGATGAA 1553  
 Qy 1261 GAGAGTGGCAAGCAGGAGGAGGAATCTGTATGAACCACTTTAGAGAACTTGCCTGAG 1320  
 Db 1554 GAGAGTGGCAAGCAGGAGGAGGAATCTGTATGAACCACTTTAGAGAACTTGCCTGAG 1613  
 Qy 1321 TCTCAGCTATTCTCTCCATGGACCTCCAGCACTTGAATCAATTTAGAGGATGCCACA 1380  
 Db 1614 TCTCAGCTATTCTCTCCATGGACCTCCAGCACTTGAATCAATTTAGAGGATGCCACA 1673  
 Qy 1381 TCTGACCTTCGATACCATG 1399  
 Db 1674 TCTGACCTTCGATACCATG 1692

## RESULT 5

ABK89295  
 ID ABK89295 standard; cDNA; 2869 BP.

XX ABK89295;

XX 21-OCT-2002 (first entry)

XX Human cDNA encoding novel serine/threonine serine kinase.

XX Human; ss; gene; serine/threonine protein kinase; inflammation; cancer;  
 KW arteriosclerosis; psoriasis; SNF kinase; transgenic; chromosome 3.

XX Homo sapiens.

XX Key Location/Qualifiers  
 FT CDS 181..2118  
 FT FT /\*tag= a  
 FT FT /product= "Serine/threonine protein kinase"

XX US2002082189-A1.

XX 27-JUN-2002.

XX 07-DEC-2000; 2000US-00731231.

XX 07-DEC-2000; 2000US-00731231.

XX (GUEG/) GUEGLER K.

XX (KETCH/) KETCHUM K. A.

XX (DFRA/) DI FRANCESCO V.

XX (BEAS/) BEASLEY E M.

XX Guegler K, Ketchum KA, Di Francesco V, Beasley EM;

XX WPI; 2002-598989/64.

XX P-FSDB; ABG31081.

XX New isolated human kinase peptide for detecting a modulator of the  
 PT peptide's expression, activity or function, that can be used to treat  
 PT disorders or disease.  
 XX  
 XX Claim 4; Fig 1; 321pp; English.  
 XX  
 CC The invention relates to an isolated kinase peptide, comprising, a novel  
 CC human serine/threonine protein kinase sequence, its allelic variant or  
 CC orthologue, where the variant is encoded by a nucleic acid molecule that  
 CC hybridises under stringent conditions to the opposite strand of kinase  
 CC cDNA or gene, or a fragment comprising 10 contiguous amino acids. Also  
 CC included are an antibody that selectively binds the kinase, a gene chip  
 CC comprising a nucleic acid (or its complement) which encodes the kinase, a  
 CC transgenic non-human animal comprising the nucleic acid, a nucleic acid  
 CC vector comprising the nucleic acid and a host cell comprising the vector.  
 CC The kinase is used to identify a modulator of the kinase or to identify  
 CC an agent that binds to the kinase, which can be used to treat a disease  
 CC or condition e.g. inflammation, cancer, arteriosclerosis and psoriasis.  
 CC The nucleic acid encoding the kinase is used to produce the kinase. A  
 CC detection agent is used to detect the kinase and an oligonucleotide is  
 CC used to detect a nucleic acid encoding the kinase in a sample. The kinase  
 CC can be used: (a) to raise antibodies against the kinase or to elicit  
 CC another immune response; (b) as a reagent (including a labeling reagent)  
 CC in assays to determine levels of a kinase (or its binding partner or  
 CC ligand) in biological fluids; and (c) as markers for tissues in which the  
 CC corresponding kinase is expressed. The kinase and antibodies against it  
 CC can be used in pharmacogenomic analysis. The kinase can be used to treat  
 CC a disorder characterised by an absence of, inappropriate, or unwanted  
 CC expression of the kinase. The kinase, homologous to SNF kinases (not  
 CC defined) is expressed by a gene located on human chromosome 3. The  
 CC present sequence is the cDNA encoding the kinase  
 XX  
 XX Sequence 2869 BP; 965 A; 570 C; 645 G; 689 T; 0 U; 0 Other;

Query Match 96.4%; Score 1397.4; DB 6; Length 2869;  
 Best Local Similarity 99.9%; Pred. No. 0;  
 Matches 1398; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 ATGCTGAAATTCAGAGGAGGAGCTAAGTGTGTGATGGATCAACAGCACTTCCACTTAT 60  
 Db 181 ATGCTGAAATTCAGAGGAGGAGCTAAGTGTGTGATGGATCAACAGCACTTCCACTTAT 240  
 Qy 61 CCAAGACCTTGATTGCAAGAGATAGCTGCTTCAACAAAACTTGGCAGTGAAGTTT 120  
 Db 241 CCAAGACCTTGATTGCAAGAGATAGCTGCTTCAACAAAACTTGGCAGTGAAGTTT 300  
 Qy 121 GGAACCTGTCTATCTGTTTCAGACAAGAAAGCCAAACGAGGAGGAAATTAAGGTACTT 180  
 Db 301 GGAACCTGTCTATCTGTTTCAGACAAGAAAGCCAAACGAGGAGGAAATTAAGGTACTT 360  
 Qy 181 AAGGAAATATCTGTTGAGAACTAAATCCAAATGAAACTGTACAGGCAATTTGGAAGCC 240  
 Db 361 AAGGAAATATCTGTTGAGAACTAAATCCAAATGAAACTGTACAGGCAATTTGGAAGCC 420  
 Qy 241 CAACCTCTCCAGCTGGACCCAGCCATGTCAGAGTTCCATGCAAGTTTGTGGAG 300  
 Db 421 CAACCTCTCCAGCTGGACCCAGCCATGTCAGAGTTCCATGCAAGTTTGTGGAG 480  
 Qy 301 CAAGATAATTTCTGCATTATCAGGAGTACTGTGAGGCGCGAGATCTGGACGATAAAAT 360  
 Db 481 CAAGATAATTTCTGCATTATCAGGAGTACTGTGAGGCGCGAGATCTGGACGATAAAAT 540  
 Qy 361 CAGGAATATAACAGCTGGAAAATCTTTCCAGAAAATCAAAATATAGATGGTTTATC 420  
 Db 541 CAGGAATATAACAGCTGGAAAATCTTTCCAGAAAATCAAAATATAGATGGTTTATC 600  
 Qy 421 CAGCTGCTGTGGAGTTGACTACATGATGATGAGAGGAGGATCTTCATCGAGACTTAAAG 480  
 Db 601 CAGCTGCTGTGGAGTTGACTACATGATGATGAGAGGAGGATCTTCATCGAGACTTAAAG 660  
 Qy 481 TCAAGAAATGATTTCTGAAAAATATCTCTTAAATTTGGAGATTTTGGAGTTTCTCGA 540



QY 241 CAATCTCTCTCCAAAGCTGGACCCACCCAGCCATTGTCAAGTTCCATGCAAGTTTGTGGAG 300  
Db 394 CAATCTCTCTCCAAAGCTGGACCCACCCAGCCATTGTCAAGTTCCATGCAAG- TTTGTGGAG 452  
QY 301 CAGATATATTTCTGCAATATACCGAGTACTGTGAGGCGCGAGATCTGGACGATAAAATT 360  
Db 453 CAGATATATTTCTGCAATATACCGAGTACTGTGAGGCGCGAGATCTGGACGATAAAATT 512  
QY 361 CAGGAATATAAACAAAGCTGGAATAATCTTTCCAGAAAATCAAAATAATAGAAATGTTTATC 420  
Db 513 CAGGAATATAAACAAAGCTGGAATAATCTTTCCAGAAAATCAAAATAATAGAAATGTTTATC 572  
QY 421 CAGCTGCTGCTGGAGTTGACTACATCATGAGGAGGAGATCTTATCGAGACTTAAAG 480  
Db 573 CAGCTGCTGCTGGAGTTGACTACATCATGAGGAGGAGATCTTATCGAGACTTAAAG 632  
QY 481 TCAAGAAATGTAATTTCTGAAAAATTAATCTCTTAAATAATCGAGATTTTGAGTTTCTCGA 540  
Db 633 TCAAGAAATGTAATTTCTGAAAAATTAATCTCTTAAATAATCGAGATTTTGAGTTTCTCGA 592  
QY 541 CTTCATATGGATCTGCTGACCTGGCCCAAACTTTAACTGGAACCTCCCATATATAGAT 600  
Db 693 CTTCATATGGATCTGCTGACCTGGCCCAAACTTTAACTGGAACCTCCCATATATAGAT 752  
QY 601 CTTGAGGCTCTGAAACACCAAGGCTATGACACAAAGTCGAGACATCTGTCACCTGGCATCG 660  
Db 753 CTTGAGGCTCTGAAACACCAAGGCTATGACACAAAGTCGAGACATCTGTCACCTGGCATCG 812  
QY 661 ATTGTTGATGAGATGCTGCTCATGATCAATGCAATCTGCTGGCTCCAATTTCTTATCAAT 720  
Db 813 ATTGTTGATGAGATGCTGCTCATGATCAATGCAATCTGCTGGCTCCAATTTCTTATCAAT 872  
QY 721 GTTTTAAAAATTTGTAAGTGACACACCTTCTCTCCCTGAGAGATATCAAAAGAACTA 780  
Db 873 GTTTTAAAAATTTGTAAGTGACACACCTTCTCTCCCTGAGAGATATCAAAAGAACTA 932  
QY 781 AATGCCATCATGAAAGCATGTTGAAACAAGATCTTCAATTAAGACCATCTGCTATCGAA 840  
Db 933 AATGCCATCATGAAAGCATGTTGAAACAAGATCTTCAATTAAGACCATCTGCTATCGAA 992  
QY 841 ATTTTAAAAATCTTACCTTGATGAGAGCTACAGAACCTTAATGTGTAGATATTAGAA 900  
Db 993 ATTTTAAAAATCTTACCTTGATGAGAGCTACAGAACCTTAATGTGTAGATATTAGAA 1052  
QY 901 ATGACTCTGGAAGACAAAAATTTGGATTGTCAGAGGAGGCTGCTCATATATAATTAATGCC 960  
Db 1053 ATGACTCTGGAAGACAAAAATTTGGATTGTCAGAGGAGGCTGCTCATATATAATTAATGCC 1112  
QY 961 ATGCAAAAAAGGATGCACTGAGAGCTCTGAGGGCACTCTCAGAACTACAGAAAATGAGC 1020  
Db 1113 ATGCAAAAAAGGATGCACTGAGAGCTCTGAGGGCACTCTCAGAACTACAGAAAATGAGC 1172  
QY 1021 CCAAGAAAGAGGATGCGCTGAGGAGCTCCAGGGGCTGTAGTGAAGACCCAGGAGCTG 1080  
Db 1173 CCAAGAAAGAGGATGCGCTGAGGAGCTCCAGGGGCTGTAGTGAAGACCCAGGAGCTG 1232  
QY 1081 AAAAGATTGTGAAGAAAAATATGAAGAAAAATAGCAACGAAATGCAAGAAATTTGAGATCT 1140  
Db 1233 AAAAGATTGTGAAGAAAAATATGAAGAAAAATAGCAACGAAATGCAAGAAATTTGAGATCT 1292  
QY 1141 CGGAATCTTACAGCTGAGTCTGTGATGTACTCCATGAAAAACACATTTAAAGGAATG 1200  
Db 1293 CGGAATCTTACAGCTGAGTCTGTGATGTACTCCATGAAAAACACATTTAAAGGAATG 1352  
QY 1201 GAAGAAAGAGGAGCAACTGAGGGAAGACTTTCTTGTTCACCCCGAGGAGGATGAA 1260  
Db 1353 GAAGAAAGAGGAGCAACTGAGGGAAGACTTTCTTGTTCACCCCGAGGAGGATGAA 1412  
QY 1261 GAGAGTGCAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1320  
Db 1413 GAGAGTGCAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1472

QY 1321 TCTCAGCCTATTCTCTCCATGAGCCTCCACGAACCTTGAATCAATTTGTAGAGGATGCCACA 1380  
Db 1473 TCTCAGCCTATTCTCTCCATGAGCCTCCACGAACCTTGAATCAATTTGTAGAGGATGCCACA 1532  
QY 1381 TCTGACCTTGGATACCATG 1399  
Db 1533 TCTGACCTTGGATACCATG 1551  
RESULT 7  
ADC99143  
ID ADC99143 standard; cdna; 2688 BP.  
XX  
AC ADC99143;  
XX  
DT 01-JAN-2004 (first entry)  
XX Human KPP cdna - SEQ ID 96.  
DE  
XX  
KW anti-HIV; anti-allergic; anti-inflammatory; antianemic; antiparkinsonian;  
nootropic; anticonvulsant; antiarteriosclerotic; antiasthmatic;  
immunosuppressive; antithyroid; cytostatic; hepatotropic; dermatological;  
antidiabetic; nephrotropic; antigout; thyromimetic; neuroprotective;  
osteopathic; antichratic; antiparasitic; antihelminthic; antipsoriatic;  
uropathic; ophthalmological; antirheumatic; haemostatic; antibacterial;  
viricide; protozoicide; fungicide; kinase; phosphatase; KPP;  
cell proliferative disorder; atherosclerosis; cirrhosis; hepatitis;  
cancer; developmental; mental retardation; neurological;  
Alzheimer's disease; Parkinson's; autoimmune; inflammatory;  
diabetes mellitus; viral; bacterial; fungal; parasitic; protozoan;  
helminthic infection; transgenic; gene therapy; human; ss; gene.  
OS Homo sapiens.  
XX  
XX WO2003033680-A2.  
FN  
XX  
PD 24-APR-2003.  
XX  
XX 17-OCT-2002; 2002WO-US033723.  
XX  
PF  
XX  
PR 19-OCT-2001; 2001US-0345474P.  
PR 02-NOV-2001; 2001US-0343910P.  
PR 13-NOV-2001; 2001US-0333098P.  
PR 16-NOV-2001; 2001US-0332424P.  
PR 30-NOV-2001; 2001US-0334288P.  
XX  
XX (INCY-) INCYTE GENOMICS INC.  
XX  
XX Bandman O, Baughn MR, Becha SD, Borowsky ML, Duggan BM;  
Emerling BM, Forsythe IJ, Gandhi AR, Gorvad AE, Griffin JA;  
Gururajan R, Hafalia AJA, Khan FA, Lal PG, Lee EA, Lee SY;  
Lindquist EA, Lu DAM, Lu Y, Marquis JP, Nguyen DB, Arvizu CS;  
Rakumar J, Recipon SA, Richardson TW, Swainkar A, Tang YT;  
Thornton MB, Tran UK, Chawla NK, Warren BA, Yang J, Rao MG, Yue H;  
Zebajadrian Y;  
XX  
XX WPI; 2003-403214/38.  
DR P-PSDB; ADC99091.  
XX  
XX New human kinases and phosphatases and polynucleotides, useful for  
diagnosing, treating or preventing autoimmune or inflammatory disorders  
(e.g. AIDS, allergy or anemia), multiple sclerosis, osteoarthritis,  
cancer or hepatitis.  
XX  
PS Claim 5; SEQ ID NO 96; 424pp; English.  
XX  
XX The invention relates to a novel isolated polypeptide which is a human  
kinase and phosphatase (KPP). The KPP polypeptides, polynucleotides,  
agonists and antagonists are useful for diagnosing, treating or  
preventing cell proliferative disorders such as atherosclerosis,  
cirrhosis, hepatitis and cancer, developmental disorders e.g. mental  
retardation, neurological disorders including Alzheimer's disease and  
Parkinson's disease, autoimmune and inflammatory disorders such as



CC Crohn's disease and diabetes mellitus and finally, viral, bacterial, fungal, parasitic, protozoan or helminthic infections. Furthermore, the CC polynucleotides encoding KPP may be useful for creating transgenic CC animals to model human disease, as well as during gene therapy CC procedures. The current sequence is that of the human KPP cDNA of the CC invention.

XX  
SQ Sequence 2688 BP; 842 A; 562 C; 615 G; 669 T; 0 U; 0 Other;  
Query Match 81.4%; Score 1179; DB 9; Length 2688;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 1179; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 ATGCTGAAATTCAGAGGAGCAGCTAAGTGTGTGAGTGATGAGTCAACAGCCATTTCCTTAT 60  
DB 251 ATGCTGAAATTCAGAGGAGCAGCTAAGTGTGTGAGTGATGAGTCAACAGCCATTTCCTTAT 310  
QY 61 CCAAGACCTTGTTCAGAGAGATACGCTTCAACAAACTTGGCAGTGGAGTTT 120  
DB 311 CCAAGACCTTGTTCAGAGAGATACGCTTCAACAAACTTGGCAGTGGAGTTT 370  
QY 121 GGAAGTGTCTATCTGTTTCAGACAAAGAAAGCCAAACGAGGAGAGAAATTAAGGTACTT 180  
DB 371 GGAAGTGTCTATCTGTTTCAGACAAAGAAAGCCAAACGAGGAGAGAAATTAAGGTACTT 430  
QY 181 AAGGAAATATCTTGGAGAACTAAATCCAAATGAAACTGTACAGCCAAATTTGAGGCC 240  
DB 431 AAGGAAATATCTTGGAGAACTAAATCCAAATGAAACTGTACAGCCAAATTTGAGGCC 490  
QY 241 CAAGTCTCTCAAGCTGGACCAACCCAGCCATGTCAGTTCCATCAAGTTCTTGGAG 300  
DB 491 CAAGTCTCTCAAGCTGGACCAACCCAGCCATGTCAGTTCCATCAAGTTCTTGGAG 550  
QY 301 CAAGATAATTTCTGCATTTACCGAGTACTGTGAGGCGCGAGATCTGACGATTAAGT 360  
DB 551 CAAGATAATTTCTGCATTTACCGAGTACTGTGAGGCGCGAGATCTGACGATTAAGT 610  
QY 361 CAGGATAATAACAGCTGGAATTTTCAGAAATCAATATAGATGTTTATC 420  
DB 611 CAGGATAATAACAGCTGGAATTTTCAGAAATCAATATAGATGTTTATC 670  
QY 421 CAGTCTGCTGGAGTTGACTACATGATGAGGAGGAGATCTTATCGAGCTTAAG 480  
DB 671 CAGTCTGCTGGAGTTGACTACATGATGAGGAGGAGATCTTATCGAGCTTAAG 730  
QY 481 TCAAGAAATGTTTCTGAAAATAATCTCTTAAATTTGGAGTTTGGAGTTCTCGA 540  
DB 731 TCAAGAAATGTTTCTGAAAATAATCTCTTAAATTTGGAGTTTGGAGTTCTCGA 790  
QY 541 CTCTAATGGGATCTGTGACCTGGCCACAACTTTAACTGGAATCTCCCATTTATAGT 600  
DB 791 CTCTAATGGGATCTGTGACCTGGCCACAACTTTAACTGGAATCTCCCATTTATAGT 850  
QY 601 CCTGAGGCTCTGAAACCAAGGCTATGACAAAGTCGACATCTGGTCACTGGCATGC 660  
DB 851 CCTGAGGCTCTGAAACCAAGGCTATGACAAAGTCGACATCTGGTCACTGGCATGC 910  
QY 661 ATTTTGTATGATGCTGCTGATGAATCATGCTGCTGCTGCTGCTGCTGCTGCTGCT 720  
DB 911 ATTTTGTATGATGCTGCTGATGAATCATGCTGCTGCTGCTGCTGCTGCTGCTGCT 970  
QY 721 GTTTTAAATTTGTTGAGGTGACACACCTTCTCTCCCTGAGAGATATCCAAAGAACTA 780  
DB 971 GTTTTAAATTTGTTGAGGTGACACACCTTCTCTCCCTGAGAGATATCCAAAGAACTA 1030  
QY 781 AATGCCATCATGGAAGAGTGTGGAACAGATCTTCAATTAAGACCATCTGCTATCGAA 840  
DB 1031 AATGCCATCATGGAAGAGTGTGGAACAGATCTTCAATTAAGACCATCTGCTATCGAA 1090  
QY 841 ATTTTAAATTCCTTACCTTGTATGAGCAGCTACAGAACCTTAATGTGTAGATATTCAGAA 900  
DB 1091 ATTTTAAATTCCTTACCTTGTATGAGCAGCTACAGAACCTTAATGTGTAGATATTCAGAA 1150

QY 901 ATGACTCTGGAAGACAAAATTTGGATTGTGAGAGGAGGCTGTATATAATTAATGCC 960  
DB 1151 ATGACTCTGGAAGACAAAATTTGGATTGTGAGAGGAGGCTGTATATAATTAATGCC 1210  
QY 961 ATGCAAAAAGGATCCACCTGCGACTCTGAGGSCACTGTGAGAGTACAGAAAATGACG 1020  
DB 1211 ATGCAAAAAGGATCCACCTGCGACTCTGAGGSCACTGTGAGAGTACAGAAAATGACG 1270  
QY 1021 CCAAGAGAAAGGATGCGGCTGAGGAAGCTCCAGCGGCTGATGAGAAAGCCAGGAGCTG 1080  
DB 1271 CCAAGAGAAAGGATGCGGCTGAGGAAGCTCCAGCGGCTGATGAGAAAGCCAGGAGCTG 1330  
QY 1081 AAAAGAGTTTGGAGAAATATGAGAAATATGAGAAATATGAGAAATATGAGAAATGAGATCT 1140  
DB 1331 AAAAGAGTTTGGAGAAATATGAGAAATATGAGAAATATGAGAAATATGAGATCT 1390  
QY 1141 CGGAACCTTTCAGCAGCTGAGTGTTCATGTTACTCCATGAA 1179  
DB 1391 CGGAACCTTTCAGCAGCTGAGTGTTCATGTTACTCCATGAA 1429  
RESULT 8  
ADE28343  
ID ADE28343 standard; cDNA; 2631 BP.  
XX  
AC ADE28343;  
XX 29-JAN-2004 (first entry)  
XX Human KPP cDNA - SEQ ID 54.  
XX kinase; phosphatase; KPP; hepatotropic; antiarteriosclerotic;  
KW antiproliferative; cytosolic; haemostatic; muscular; cerebroprotective;  
KW nontropic; ophthalmological; anticonvulsant; vasotropic; neuroprotective;  
KW antiparkinsonian; antiasthmatic; antianemic; antiasthmatic;  
KW antidiabetic; antiinflammatory; osteopathic; antiarthritic;  
KW antirheumatic; dermatological; virucide; antibacterial; fungicide;  
KW antiparasitic; protozoic; antihelminthic; antitumor; cardiovascular;  
KW antiarteriosclerotic; immunosuppressive; cell proliferative; cirrhosis;  
KW hepatitis; arteriosclerosis; psoriasis; primary thrombocytopenia; cancer;  
KW developmental; renal tubular acidosis; Becker's muscular dystrophy;  
KW gonadal dysgenesis; hypothyroidism; seizure; neurological;  
KW Pick's disease; cataract; epilepsy; ischaemic cerebrovascular; stroke;  
KW Alzheimer's; Parkinson's; dementia; autoimmune; inflammatory; AIDS;  
KW allergy; anaemia; asthma; diabetes mellitus; bronchitis; osteoporosis;  
KW osteoarthritis; rheumatoid arthritis; contact dermatitis; gout;  
KW lipid disorder; cholestasis; Gaucher's; diabetes; atherosclerosis; liver;  
KW viral; bacterial; fungal; parasitic; protozoan; helminthic infection;  
KW trauma; gene therapy; human; ss; gene.  
XX Homo sapiens.  
XX WO2003080805-A2.  
XX 02-OCT-2003.  
XX 18-MAR-2003; 2003WO-US008715.  
XX 19-MAR-2002; 2002US-0366088P.  
XX 29-MAR-2002; 2002US-0369248P.  
XX (INCY) INCYTE CORP.  
XX Chien D, Jin P, Hawkins PR, Baughn MR, Becha SD, Chang H;  
XX Ding L, Elliott VS, Emerling BM, Gandhi AR, Gietzen KJ, Griffin JA;  
XX Gururajan R, Hafalia AJA, Ison CH, Kable AE, Khare R, Lee SY;  
XX Lee EA, Lu Y, Marquis JP, Lehr-Wason PM, Ramkumar J, Richardson TW;  
XX Swarnakar A, Tran UK, Chawla NK, Yao MG, Yue H, Bhatia U;  
XX Burrill JD, Lee S, Blake JJ, Ho A, Zheng W;  
XX WPI; 2004-011523/01.  
XX P-PSDB; ADE28291.



PT New human kinases and phosphatases, and polynucleotides encoding them,  
 PT useful for treating, preventing or diagnosing e.g. cell proliferative  
 PT disorders, inflammatory, autoimmune, viral, bacterial, parasitic or  
 XX fungal diseases.

PS Claim 5; SEQ ID NO 54; 340pp; English.

XX  
 CC The invention relates to a novel isolated kinase and phosphatase (KPP)  
 CC polypeptide. The polypeptide of the invention demonstrates hepatocytic,  
 CC antiarteriosclerotic, antiparasitic, cytostatic, haemostatic, muscular,  
 CC cerebroprotective, neurotropic, ophthalmological, anticonvulsant,  
 CC vasoprotective, antiparkinsonian, antidiabetic,  
 CC antiasthmatic, antiinflammatory, antiparasitic, osteopathic,  
 CC antiarthritic, antirheumatic, dermatological, virucide, antibacterial,  
 CC fungicide, antiparasitic, protozoacide, antihelminthic, antitumor,  
 CC cardiovascular, antiarteriosclerotic and immunosuppressive activities.  
 CC The KPP polypeptides may be useful for diagnosing, treating or preventing  
 CC cell proliferative disorders including cirrhosis, hepatitis,  
 CC arteriosclerosis, psoriasis, primary thrombocytopenia and cancer,  
 CC developmental disorders such as renal tubular acidosis, Becker's muscular  
 CC dystrophy, gonadal dysgenesis, hypothyroidism or seizures, neurological  
 CC disorders e.g. Pick's disease, cataract, epilepsy, ischaemic  
 CC cerebrovascular disease, stroke, Alzheimer's disease, Parkinson's disease  
 CC or dementia, autoimmune or inflammatory disorders including AIDS,  
 CC allergies, anaemia, asthma, diabetes mellitus, bronchitis, osteoporosis,  
 CC osteoarthritis, rheumatoid arthritis, contact dermatitis or gout and  
 CC lipid disorders such as cholestasis, Gaucher's disease, diabetes,  
 CC atherosclerosis or liver disease, as well as viral, bacterial, fungal,  
 CC parasitic, protozoan or helminthic infections and trauma. Furthermore,  
 CC the polypeptide may be utilised during gene therapy procedures. The  
 CC current sequence is that of the human KPP cDNA of the invention.

XX SQ Sequence 2631 BP; 833 A; 539 C; 619 G; 640 T; 0 U; 0 Other;

Query Match 73.48; Score 1063; DB 10; Length 2631;  
 Best Local Similarity 100.0%; Pred. No. 2.9e-300;  
 Matches 1063; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 337 GCGGAGATCTGGACGATAAAATTCAGGAATATAACAGCTGGAAAAATCTTTCAGAA 396  
 DB 421 GCGGAGATCTGGACGATAAAATTCAGGAATATAACAGCTGGAAAAATCTTTCAGAA 480  
 QY 397 AATCAATATAGATGTTTATCCAGCTCTCTCTGGAGTTGACTACATGATGAGAG 456  
 DB 481 AATCAATATAGATGTTTATCCAGCTCTCTCTGGAGTTGACTACATGATGAGAG 540  
 QY 457 AGGATATCTTCATCGAGACTTAAAGTCAAAAGATGTTTCTGAAAAATATCTCTTAA 516  
 DB 541 AGGATATCTTCATCGAGACTTAAAGTCAAAAGATGTTTCTGAAAAATATCTCTTAA 600  
 QY 517 ATTGGAGATTTGGAGTTTCTGACTTCTAATGGGATCTGTGACCTGGCCCAACTTTA 576  
 DB 601 ATTGGAGATTTGGAGTTTCTGACTTCTAATGGGATCTGTGACCTGGCCCAACTTTA 660  
 QY 577 ACTGGAACTCCCATATATAGTCTCTGAGCTCTGAAACCAACAGGCTATGACACAAAG 636  
 DB 661 ACTGGAACTCCCATATATAGTCTCTGAGCTCTGAAACCAACAGGCTATGACACAAAG 720  
 QY 637 TCGGACATCTGGTCACTGGCATCTTTGTATGAGATGTGCTGCAATGATCATGCTTC 696  
 DB 721 TCGGACATCTGGTCACTGGCATCTTTGTATGAGATGTGCTGCAATGATCATGCTTC 780  
 QY 697 GCTGGCTCCAAATTTCTATCTATGTTTAAATTTTGAAGGTGACACACTTCTCTC 756  
 DB 781 GCTGGCTCCAAATTTCTATCTATGTTTAAATTTTGAAGGTGACACACTTCTCTC 840  
 QY 757 CCTGAGAGATATCCAAAAGAACTAAATGCGCATCATGGAAGCATGTTGAACAAGATCT 816  
 DB 841 CCTGAGAGATATCCAAAAGAACTAAATGCGCATCATGGAAGCATGTTGAACAAGATCT 900  
 QY 817 TCATTAAAGCACTCTGCTATCGAAATTTTAAATTCCTTACCTGATGAGCAGCTACG 876  
 DB 901 TCATTAAAGCACTCTGCTATCGAAATTTTAAATTCCTTACCTGATGAGCAGCTACG 960

QY 877 AACCTAATGTGTAGATATTTCAGAAATGACTCTGGAAGACAAAATTTGGATTGTGAGAAG 936  
 DB 961 AACCTAATGTGTAGATATTTCAGAAATGACTCTGGAAGACAAAATTTGGATTGTGAGAAG 1020  
 QY 937 GAGGCTGTCTATATAATTAATGCCATGCAAAAAGGATCCACCTGCAGACTCTGAGGSCA 996  
 DB 1021 GAGGCTGTCTATATAATTAATGCCATGCAAAAAGGATCCACCTGCAGACTCTGAGGSCA 1080  
 QY 997 CTGTGAGAGTACAGAAATACGCCAAGAGAAAGGATCCGGCTGAGGAAGCTCCAGGCG 1056  
 DB 1081 CTGTGAGAGTACAGAAATACGCCAAGAGAAAGGATCCGGCTGAGGAAGCTCCAGGCG 1140  
 QY 1057 GCTGATGAGAGCCAGAGCTGAAAAGATTGTGGAAGAAAATATGAAGAAAATAGC 1116  
 DB 1141 GCTGATGAGAGCCAGAGCTGAAAAGATTGTGGAAGAAAATATGAAGAAAATAGC 1200  
 QY 1117 AAACCAATGCAAGATTCAGATCTCGGAACCTTCAGCACTGAGTGTGTGATCTACTCCAT 1176  
 DB 1201 AAACCAATGCAAGATTCAGATCTCGGAACCTTCAGCACTGAGTGTGTGATCTACTCCAT 1260  
 QY 1177 GAAAGAACATTTAAAGGAATGGAAGAAAGAGGAGAACCTGAGGGAAGACTTTCT 1236  
 DB 1261 GAAAGAACATTTAAAGGAATGGAAGAAAGAGGAGAACCTGAGGGAAGACTTTCT 1320  
 QY 1237 TGTTCACCCAGGACGAGATGAAGAGAGTGGCAAGGAGGAGGAGGAGGAGGAGGAGG 1296  
 DB 1321 TGTTCACCCAGGACGAGATGAAGAGAGTGGCAAGGAGGAGGAGGAGGAGGAGGAGG 1380  
 QY 1297 CCACTTTAGAGAACCTGCTGAGTCTCAGCCTATTCCTTCCATGAGCCTCCAGAACTT 1356  
 DB 1381 CCACTTTAGAGAACCTGCTGAGTCTCAGCCTATTCCTTCCATGAGCCTCCAGAACTT 1440  
 QY 1357 GAATCAATTTAGAGGATGCCACATCTGACCTTGTATACCATG 1399  
 DB 1441 GAATCAATTTAGAGGATGCCACATCTGACCTTGTATACCATG 1483

RESULT 9  
 ADC99142  
 ID ADC99142 standard; cDNA; 2483 BP.  
 XX  
 AC ADC99142;  
 XX  
 DT 01-JAN-2004 (first entry)  
 XX  
 DE Human KPP cDNA - SEQ ID 95.  
 XX  
 KW anti-HIV; antiallergic; antiinflammatory; antianemic; antiparkinsonian;  
 KW neurotic; anticonvulsant; antiarteriosclerotic; antidiabetic;  
 KW immunosuppressive; antihypertensive; cytostatic; hepatocytic; dermatological;  
 KW antidiabetic; nephroprotective; antitumor; thymostatic; neuroprotective;  
 KW osteopathic; antiparasitic; antihelminthic; antiparasitic; antiparasitic;  
 KW uropathic; ophthalmological; antirheumatic; haemostatic; antibacterial;  
 KW virucide; protozoacide; fungicide; kinase; phosphatase; KPP;  
 KW cell proliferative disorder; atherosclerosis; cirrhosis; hepatitis;  
 KW cancer; developmental disorder; mental retardation; neurological;  
 KW Alzheimer's disease; Parkinson's; autoimmune; inflammatory; Crohn's;  
 KW diabetes mellitus; viral; bacterial; fungal; parasitic; protozoan;  
 KW helminthic infection; transgenic; gene therapy; human; ss; gene.  
 OS Homo sapiens.  
 XX  
 XX WO200303680-A2.  
 XX  
 PD 24-APR-2003.  
 XX  
 XX 17-OCT-2002; 2002WO-US033723.  
 XX  
 PR 19-OCT-2001; 2001US-0345474P.  
 PR 02-NOV-2001; 2001US-0343910P.  
 PR 13-NOV-2001; 2001US-0333098P.  
 PR 16-NOV-2001; 2001US-0332424P.

```
PR 30-NOV-2001; 2001US-0334288P.
XX (INCY-) INCYTE GENOMICS INC.
XX
XX Bandman O, Baughn MR, Becha SD, Borowsky ML, Duggan BM;
XX Emerling BV, Forsythe IJ, Gandhi AR, Gorvad AE, Griffin JB;
XX Gururajan R, Hafalia AJA, Khan FA, Lal PG, Lee EA, Lee SY;
XX Lindquist EA, Lu DAM, Lu Y, Marquis JP, Nguyen DB, Arvizu CS;
XX Ramkumar J, Recipon SA, Richardson TW, Swarnakar A, Tang YT;
XX Thornton MB, Tran UK, Chawla NK, Warren BA, Yang J, Yac MG, Yue H;
XX Zebbarjadian Y;
XX WPI: 2003-403214/38.
XX P-PSDB; ADC99090.
XX
XX New human kinases and phosphatases and polynucleotides, useful for
XX diagnosing, treating or preventing autoimmune or inflammatory disorders
XX (e.g. AIDS, allergy or anemia), multiple sclerosis, osteoarthritis,
XX cancer or hepatitis.
XX
XX Claim 5; SEQ ID NO 95; 424pp; English.
XX
XX The invention relates to a novel isolated polypeptide which is a human
XX kinase and phosphatase (KPP). The KPP polypeptides, polynucleotides,
XX agonists and antagonists are useful for diagnosing, treating or
XX preventing cell proliferative disorders such as atherosclerosis,
XX cirrhosis, hepatitis and cancer, developmental disorders e.g. mental
XX retardation, neurological disorders including Alzheimer's disease and
XX Parkinson's disease, autoimmune and inflammatory disorders such as
XX Crohn's disease and diabetes mellitus and finally, viral, bacterial,
XX fungal, parasitic, protozoan or helminthic infections. Furthermore, the
XX polynucleotides encoding KPP may be useful for creating transgenic
XX animals to model human disease, as well as during gene therapy
XX procedures. The current sequence is that of the human KPP cDNA of the
XX invention.
XX
XX SQ Sequence 2483 BP; 783 A; 515 C; 597 G; 588 T; 0 U; 0 Other;
XX
XX Query Match 52.4%; Score 759; DB 9; Length 2483;
XX Best Local Similarity 77.5%; Pred No. 3.3e-211;
XX Matches 1084; Conservative 0; Mismatches 0; Indels 315; Gaps 1;
XX
XX 1 ATGCTGAAATCCCAAGAGGAGGAGCTGAGTGTGAGTGGATCAACAGCCATTCCACTTAT 60
XX 251 ATGCTGAAATCCCAAGAGGAGGAGCTGAGTGTGAGTGGATCAACAGCCATTCCACTTAT 310
XX
XX 61 CCAAGACCTTGATTGCAAGAGATAGTCTTCAACAAAACCTTGGCAGTGGAGTTT 120
XX 311 CCAAGACCTTGATTGCAAGAGATAGTCTTCAACAAAACCTTGGCAGTGGAGTTT 370
XX
XX 121 GGAATCTGTATCTGGTTTTCAGACAAGAAAGCCAAACGAGGAGGAGGAATTAAGGTACTT 180
XX 371 GGAATCTGTATCTGGTTTTCAGACAAGAAAGCCAAACGAGGAGGAGGAATTAAGGTACTT 430
XX
XX 181 AAGAAATATCTGTGAGAACTAAATCCAAATCAAACTGTACAGGCCAATTTGGAGCC 240
XX 431 AAGAAATATCTGTGAGAACTAAATCCAAATCAAACTGTACAGGCCAATTTGGAGCC 490
XX
XX 241 CAAGTCTCTCCAAAGCTGGACCCAGCCAGCAATCTCAAGTTCCATGCAAGTTTGTGGAG 300
XX 491 CAAGTCTCTCCAAAGCTGGACCCAGCCAGCAATCTCAAGTTCCATGCAAGTTTGTGGAG 550
XX
XX 301 CAAGATAATTTCTGCATTATCAAGGATATCTGTGAGGGCCGAGATCTGGACGATAAATT 360
XX 551 CAAGATAATTTCTGCATTATCAAGGATATCTGTGAGGGCCGAGATCTGGACGATAAATT 610
XX
XX 361 CAGGAATATAACAGCTGGAAATCTTTTCCGAAATCAATAATAGATGTTTATC 420
XX 611 CAGGAATATAACAGCTGGAAATCTTTTCCGAAATCAATAATAGATGTTTATC 670
XX
XX 421 CAGCTGCTGCTGGAGTTGATACATCATGATGAGAGGAGGATCTTCAATCGAGACTTAAG 480
XX 671 CAGCTGCTGCTGGAGTTGATACATCATGATGAGAGGAGGATCTTCAATCGAGACTTAAG 730
```

RESULT 10

ABL34984  
ID ABL34984 standard; cDNA; 2538 BP.

XX ABL34984;

XX AC

XX 04-APR-2002 (first entry)

DT

XX DE Rat cDNA isolated from skin cells SEQ ID NO: 511.  
 XX KW Human; rat; mouse; skin cell; skin wound; cancer; growth defect;  
 KW developmental defect; inflammatory disease; dermatological; vulnary;  
 KW immunomodulator; anti-inflammatory; cytostatic; neuroprotective; gene;  
 KW ss.  
 XX OS Rattus sp.  
 XX PN WO2001:90357-A1.  
 XX XX 29-NOV-2001.  
 XX PD 24-MAY-2001; 2001WO-NZ000099.  
 XX XX 24-MAY-2000; 2000US-0206650P.  
 XX PR 25-JUL-2000; 2000US-0221232P.  
 XX XX (GENE-) GENESIS RES & DEV CORP LTD.  
 XX PA Watson JD, Strachan L, Sleeman M, Onrust R, Murison JG;  
 PI Kumble KD;  
 PI WPI; 2002-122020/16.  
 XX DR New polynucleotides and polypeptides encoded by the polynucleotides  
 PT isolated from skin cells, useful for treating skin wounds, cancers,  
 PT growth and developmental defects, inflammatory diseases, or for  
 PT modulating immune responses.  
 XX CC Claim 1; Page 316-317; 466pp; English.  
 XX CC The present invention provides the protein and coding sequences of cDNAs  
 CC isolated from human, murine and rat skin cell libraries. The sequences  
 CC can be used in the development of therapeutic agents useful in the  
 CC treatment of skin diseases, including skin wounds, cancer, growth  
 CC defects, developmental defects and inflammatory diseases. The proteins  
 CC have important roles in the induction of hair growth, cell proliferation  
 CC and cell-cell interaction, in maintaining tissue integrity, in wound  
 CC healing and in modulating immune responses. The present sequence is a  
 CC cDNA of the invention  
 XX SQ Sequence 2538 BP; 697 A; 558 C; 636 G; 647 T; 0 U; 0 Other;  
 Query Match 46.2%; Score 669.6; DB 6; Length 2538;  
 Best Local Similarity 82.2%; Pred. No. 5.1e-185;  
 Matches 797; Conservative 0; Mismatches 164; Indels 9; Gaps 2;  
 Qy 1 ATGCTGAAATCCAGAGGAGCTAGCTGCTGAGTGGATCAACAGCCATTTCCACTTAT 60  
 Db 457 ATGCTGAAATCCAGAGGAGCTAGCTGCTGAGTGGATCAACAGCCATTTCCACTTAT 510  
 Qy 61 CCAAGACCTGATTCGAGAGGATGCTGCTTCAACAAAATTCGAGTGGAGCTT 120  
 Db 511 CCACCCACTGATTCGAGAGGATGCTGCTTCAACAAAATTCGAGTGGAGCTT 570  
 Qy 121 GGAACCTGCTATCTGCTTTCAGACCAAGAGGAGGAGGAGGAGGAGGAGTACTT 180  
 Db 571 GGAACCTGCTATCTGCTTTCAGACCAAGAGGAGGAGGAGGAGGAGGAGTACTT 630  
 Qy 181 AAGAAATATCTGTTGAGAGGATTAATCCAAATGAAATGAACTGACAGGCAATTTGGAGCC 240  
 Db 631 AAGAAATATCTGTTGAGAGGATTAATCCAAATGAAATGAACTGACAGGCAATTTGGAGCC 690  
 Qy 241 CAAGCTCTCTCCAGCTGGACCCAGCCAGCTTGTCAAGTTTCCATGCAAGTTTGTGGAG 300  
 Db 691 CAGCTCTCTCCAGCTTAACCATCCGCGCATCGTCAGGTTCCACGACGTTTCATGGAG 750  
 Qy 301 CAAGATTAATTTCTGCATATCA CGAGTACTGTGAGGCGCGAGATCTGGACGATAAAAT 360  
 Db 751 CAGGATCAATTTTGCATATCA CGAGTACTGTGAGGCGCGAGATCTGGACGATAAAATC 810

361 CAGGAATATAAACAAGCTGGAAAAATCTTTCAGAAAAATCAAAATATAGATGGTTTATC 420  
 Db 811 CAGGAATATAAACAAGCTGGAAAAATCTTTCAGAAAAATCAAAATATAGATGGTTTATC 870  
 Qy 421 CAGCTGCTGGGAGTTGACTACATGATGAGAGGAGGAGTACTTCATCGAGACTTAAAG 480  
 Db 871 CAGTTGCTGCTGGGAGTTGATTACATGATGAGAGGAGGAGTACTTCATCGAGACTTAAAG 930  
 Qy 481 TCAAGAATATATTTCTGAAAAATATCTCCCTAAAAATTTGGAGATTTTGGAGTTTCTCGA 540  
 Db 931 TCAAGAATATATTTCTGAAAAATATCTCCCTAAAAATTTGGAGATTTTGGAGTTTCTCGG 990  
 Qy 541 CTCTTAATGGGATCTGTGACTGGCCCAACTTTTAAGTGAACCTCCCATATATATGAGT 600  
 Db 991 CTGCTGATGGGTTTCAATGAGCTGGCTACAACTTAAACGGGACTTCCCATATATATGAGT 1050  
 Qy 601 CCTGAGGCTCTGAAACACCAAGGCTATGACACAAAGTCCGACATCTGCTCACTGGCATGC 660  
 Db 1051 CCAGAGGCCCTGAGCACCACCAAGGCTGTGATGCCAAGTCTGACATCTGCTCACTGGCATGC 1110  
 Qy 661 ATTTTGTATGAGATGCTGTGATGATGATGATGATGATGATGATGATGATGATGATGAT 720  
 Db 1111 ATTTTATAGAGATGCTGTGATGATGATGATGATGATGATGATGATGATGATGATGAT 1170  
 Qy 721 GTTTTAAAAATTTGTAAGGTGACACACCTTCTCTCCCTGAGAGATATCCAAAAGAACTA 780  
 Db 1171 GTTTTGAATATCTGTTGAAGGCAACACGCTTCACTCCCTGACATACCCACGAGAACTA 1230  
 Qy 781 AATGCCATCATGGAAGCATGTTGAACAAAGATCTTCAATTAAGACCATCTGCTATCGAA 840  
 Db 1231 AACATCATCATGTCACGCATGTTGAACAAAGATCTTCCCTCGAGACCGTCCGCTCGCAGAG 1290  
 Qy 841 ATTTTAAAAATCCCTTACCTTGATGAGCAGCTACAGAACCTTAATGTGTAGATATTCAGAA 900  
 Db 1291 ATTTTAAAGGCCCTTATGTGAAGAGTGCCTTCAGACGCTGATGATTAACATCCAGAG 1350  
 Qy 901 ATGACTCTGGAGACAA---AAATTTGGATTTGTGAGAGGAGGCTGCTCATATTAATTAAT 957  
 Db 1351 GCGACACTGGAGACAAAGAGGAGTCTCAGCTGTGAGAGGAGGCTGCCCATGCAGTTAAC 1410  
 Qy 958 GCCATGCGAA 967  
 Db 1411 GCCATAAGA 1420  
 RESULT 11  
 AD51181  
 ID AD51181 standard; DNA; 1071 BP.  
 XX AC AD51181;  
 XX DT 18-DEC-2003 (first entry)  
 XX XX Human cell-cycle related protein coding sequence, SEQ ID 7.  
 DE Human; cytostatic; cell-cycle related protein; nuclear export;  
 KW nuclear-cytoplasm transport; cytotoxic; cell-cycle control;  
 KW immunological disease; neurological disease; cancer; gene; ds.  
 XX OS Homo sapiens.  
 XX PH Location/Qualifiers  
 FT 1..1071  
 FT /tag= a  
 FT /partial  
 FT /product= "Cell-cycle related protein"  
 FT /note= "No start or stop codon given"  
 XX XX JP2003144168-A.  
 XX XX 20-MAY-2003.  
 XX XX 14-NOV-2001; 2001JP-00349158.  
 XX PF



```
PR 06-SEP-2000; 2000US-0230437P.
PR 06-SEP-2000; 2000US-0230438P.
PR 08-SEP-2000; 2000US-0231242P.
PR 08-SEP-2000; 2000US-0231243P.
PR 08-SEP-2000; 2000US-0231244P.
PR 08-SEP-2000; 2000US-0231413P.
PR 08-SEP-2000; 2000US-0231414P.
PR 08-SEP-2000; 2000US-0232080P.
PR 08-SEP-2000; 2000US-0232081P.
PR 12-SEP-2000; 2000US-0231968P.
PR 14-SEP-2000; 2000US-0232397P.
PR 14-SEP-2000; 2000US-0232398P.
PR 14-SEP-2000; 2000US-0232399P.
PR 14-SEP-2000; 2000US-0232400P.
PR 14-SEP-2000; 2000US-0232401P.
PR 14-SEP-2000; 2000US-0233063P.
PR 14-SEP-2000; 2000US-0233064P.
PR 14-SEP-2000; 2000US-0233065P.
PR 21-SEP-2000; 2000US-0234223P.
PR 21-SEP-2000; 2000US-0234274P.
PR 25-SEP-2000; 2000US-0234997P.
PR 25-SEP-2000; 2000US-0234998P.
PR 26-SEP-2000; 2000US-0235484P.
PR 27-SEP-2000; 2000US-0235834P.
PR 27-SEP-2000; 2000US-0235836P.
PR 29-SEP-2000; 2000US-0236377P.
PR 29-SEP-2000; 2000US-0236367P.
PR 29-SEP-2000; 2000US-0236368P.
PR 29-SEP-2000; 2000US-0236369P.
PR 29-SEP-2000; 2000US-0236370P.
PR 02-OCT-2000; 2000US-0236802P.
PR 02-OCT-2000; 2000US-0237037P.
PR 02-OCT-2000; 2000US-0237038P.
PR 02-OCT-2000; 2000US-0237039P.
PR 02-OCT-2000; 2000US-0237040P.
PR 13-OCT-2000; 2000US-0239935P.
PR 13-OCT-2000; 2000US-0239937P.
PR 20-OCT-2000; 2000US-0240960P.
PR 20-OCT-2000; 2000US-0241221P.
PR 20-OCT-2000; 2000US-0241785P.
PR 20-OCT-2000; 2000US-0241786P.
PR 20-OCT-2000; 2000US-0241787P.
PR 20-OCT-2000; 2000US-0241808P.
PR 20-OCT-2000; 2000US-0241809P.
PR 20-OCT-2000; 2000US-0241826P.
PR 01-NOV-2000; 2000US-0244617P.
PR 08-NOV-2000; 2000US-0246474P.
PR 08-NOV-2000; 2000US-0246475P.
PR 08-NOV-2000; 2000US-0246476P.
PR 08-NOV-2000; 2000US-0246477P.
PR 08-NOV-2000; 2000US-0246478P.
PR 08-NOV-2000; 2000US-0246523P.
PR 08-NOV-2000; 2000US-0246524P.
PR 08-NOV-2000; 2000US-0246525P.
PR 08-NOV-2000; 2000US-0246526P.
PR 08-NOV-2000; 2000US-0246527P.
PR 08-NOV-2000; 2000US-0246528P.
PR 08-NOV-2000; 2000US-0246532P.
PR 08-NOV-2000; 2000US-0246533P.
PR 08-NOV-2000; 2000US-0246611P.
PR 08-NOV-2000; 2000US-0246612P.
PR 17-NOV-2000; 2000US-0249211P.
PR 17-NOV-2000; 2000US-0249212P.
PR 17-NOV-2000; 2000US-0249213P.
PR 17-NOV-2000; 2000US-0249214P.
PR 17-NOV-2000; 2000US-0249215P.
PR 17-NOV-2000; 2000US-0249216P.
PR 17-NOV-2000; 2000US-0249217P.
PR 17-NOV-2000; 2000US-0230437P.
PR 17-NOV-2000; 2000US-0230438P.
PR 17-NOV-2000; 2000US-0249242P.
PR 17-NOV-2000; 2000US-0249243P.
PR 17-NOV-2000; 2000US-0249244P.
PR 17-NOV-2000; 2000US-0249245P.
PR 17-NOV-2000; 2000US-0249246P.
PR 17-NOV-2000; 2000US-0249247P.
PR 17-NOV-2000; 2000US-0249248P.
PR 17-NOV-2000; 2000US-0249249P.
PR 17-NOV-2000; 2000US-0249250P.
PR 17-NOV-2000; 2000US-0249251P.
PR 17-NOV-2000; 2000US-0249252P.
PR 17-NOV-2000; 2000US-0249253P.
PR 17-NOV-2000; 2000US-0249254P.
PR 17-NOV-2000; 2000US-0249255P.
PR 17-NOV-2000; 2000US-0249256P.
PR 17-NOV-2000; 2000US-0249257P.
PR 17-NOV-2000; 2000US-0249258P.
PR 17-NOV-2000; 2000US-0249259P.
PR 17-NOV-2000; 2000US-0249260P.
PR 17-NOV-2000; 2000US-0249261P.
PR 17-NOV-2000; 2000US-0249262P.
PR 17-NOV-2000; 2000US-0249263P.
PR 17-NOV-2000; 2000US-0249264P.
PR 17-NOV-2000; 2000US-0249265P.
PR 17-NOV-2000; 2000US-0249266P.
PR 17-NOV-2000; 2000US-0249267P.
PR 17-NOV-2000; 2000US-0249268P.
PR 17-NOV-2000; 2000US-0249269P.
PR 17-NOV-2000; 2000US-0249270P.
PR 17-NOV-2000; 2000US-0249271P.
PR 17-NOV-2000; 2000US-0249272P.
PR 17-NOV-2000; 2000US-0249273P.
PR 17-NOV-2000; 2000US-0249274P.
PR 17-NOV-2000; 2000US-0249275P.
PR 17-NOV-2000; 2000US-0249276P.
PR 17-NOV-2000; 2000US-0249277P.
PR 17-NOV-2000; 2000US-0249278P.
PR 17-NOV-2000; 2000US-0249279P.
PR 17-NOV-2000; 2000US-0249280P.
PR 17-NOV-2000; 2000US-0249281P.
PR 17-NOV-2000; 2000US-0249282P.
PR 17-NOV-2000; 2000US-0249283P.
PR 17-NOV-2000; 2000US-0249284P.
PR 17-NOV-2000; 2000US-0249285P.
PR 17-NOV-2000; 2000US-0249286P.
PR 17-NOV-2000; 2000US-0249287P.
PR 17-NOV-2000; 2000US-0249288P.
PR 17-NOV-2000; 2000US-0249289P.
PR 17-NOV-2000; 2000US-0249290P.
PR 17-NOV-2000; 2000US-0249291P.
PR 17-NOV-2000; 2000US-0249292P.
PR 17-NOV-2000; 2000US-0249293P.
PR 17-NOV-2000; 2000US-0249294P.
PR 17-NOV-2000; 2000US-0249295P.
PR 17-NOV-2000; 2000US-0249296P.
PR 17-NOV-2000; 2000US-0249297P.
PR 17-NOV-2000; 2000US-0249298P.
PR 17-NOV-2000; 2000US-0249299P.
PR 17-NOV-2000; 2000US-0249300P.
PR 17-NOV-2000; 2000US-0250160P.
PR 01-DEC-2000; 2000US-0250391P.
PR 01-DEC-2000; 2000US-0250392P.
PR 05-DEC-2000; 2000US-0251030P.
PR 05-DEC-2000; 2000US-0251988P.
PR 05-DEC-2000; 2000US-0256719P.
PR 06-DEC-2000; 2000US-0251479P.
PR 08-DEC-2000; 2000US-0251856P.
PR 08-DEC-2000; 2000US-0251868P.
PR 08-DEC-2000; 2000US-0251869P.
PR 08-DEC-2000; 2000US-0251989P.
PR 08-DEC-2000; 2000US-0251990P.
PR 11-DEC-2000; 2000US-0254097P.
PR 05-JAN-2001; 2001US-0259678P.
XX
PA (HUMA-) HUMAN GENOME SCI INC.
XX
XX Rosen CA, Barash SC, Ruben SM;
PI WPI; 2001-581633/65.
DR P-PSDB; AAU87140.
XX
XX New isolated nucleic acid encoding a protein for diagnosing, preventing,
PT treating or ameliorating medical conditions and used as food additives or
PT preservatives.
XX
XX Claim 1; SEQ ID NO 60; 837pp; English.
XX
XX The invention describes an isolated nucleic acid molecule (I) encoding a
CC novel central nervous system protein. (I) and polypeptides (II) encoded
CC by (I), are used to treat a medical conditions and in diagnosis of a
CC pathological condition. Disorders which are diagnosed or treated include
CC autoimmune diseases e.g. rheumatoid arthritis, hyperproliferative
CC disorders e.g. neoplasms of the breast or liver, cardiovascular disorders
CC e.g. cardiac arrest, cerebrovascular disorders e.g. cerebral ischaemia,
CC angiogenesis, nervous system disorders e.g. Alzheimer's disease and
CC amyotrophic lateral sclerosis, infections caused by bacteria, viruses
CC e.g. Acquired immunodeficiency virus (AIDS) and fungi, ocular disorders
CC e.g. corneal infection, gastrointestinal disorders e.g. dysphagia,
CC adenocarcinomas and irritable bowel syndrome, reproductive system
CC disorders e.g. testicular feminisation, endocrine disorders e.g. diabetes
CC and pituitary dwarfism, cancers and disorders at the cellular level e.g.
CC leukaemia, disorders involving neovascularisation e.g. malignancies,
CC respiratory disorders e.g. nonallergic rhinitis, renal disorders e.g.
CC acute kidney failure and blood related disorders e.g. myocardial
CC infarction. The polypeptides can also be used to aid wound healing and
CC epithelial cell proliferation, to prevent skin aging due to sunburn, to
CC maintain organs before transplantation, for supporting cell culture of
CC primary tissues, to regenerate tissues and in chemotaxis. The
CC polypeptides can also be used as a food additive or preservative to
CC increase or decrease storage capabilities, fat content, lipid, protein,
Query Match 28.7%; Score 415.8; DB 4; Length 451;
Best Local Similarity 98.2%; Pred. NO. 4.5e-11;
Matches 439; Conservative 0; Mismatches 6; Indels 2; Gaps 2;
QY 738 AGGTGACACACCTTCTCTCCCTGAGAGATATCCAAAGAACTAAATGCATCGGAAG 797
Db 2 AGGTGACACACCTTCTCTCCCTGAGAGATATCCAAAGAACTAAATGCATCGGAAG 61
QY 798 CATGTTGAACAGAGATCCCTTCATTAAAGACCATCTGCTATCGAAATTTTAAATCCCTTA 857
Db 62 CATGTTGAACAGAGATCCCTTCATTAAAGACCATCTGCTATCGAAATTTTAAATCCCTTA 121
QY 858 CCTTCATGAGCAGCTACAGAACCTAATGTGTAGATATTCAGAAATGCTCTGGAGACAA 917
PR 17-NOV-2000; 2000US-0249217P.
```

```
Db 122 CCTTGATGAGCAGCTACAGAACCTAATGTGTAGATATTTCAGAAATGACTCTGGAAGACAA 181
Qy 918 AAATTTGGATTGTCAGAGAGGCTGCTCATATAATTAATGCCATGCAAAAAGGATCCA 977
Db 182 AAATTTGGATTGTCAGAGAGGCTGCTCATATAATTAATGCCATGCAAAAAGGATCCA 241
Qy 978 CCTGAGACTCTGAGGGCACTGTTCAGAAAGTACAGAAATGACGCAAGAGAAAGGATGCG 1037
Db 242 CCTGAGACTCTGAGGGCACTGTTCAGAAAGTACAGAAATGACGCAAGAGAAAGGATGCG 301
Qy 1038 GCTGAGAGTCTCAGG-CGGCTGATGAGAAAGCCAGGAGCTGAAAAAGATTGTGAAG 1096
Db 302 GCTGAGAGTCTCAGGCGGCTGATGAGAAAGCCAGGAGCTGAAAAAGATTGTGAAG 361
Qy 1097 AAAAATATGAAGAAATAGCAACCAATGCAAGAAATGAGATCTCGGAACCTTTCAGCAGC 1156
Db 362 ANAATNTGAGAAATAGCAACCAATGCAAGAAATGAGATCTCGGAACCTTTCAGCAGC 421
Qy 1157 TGAGTGTG-ATGTAATCCATGAAAA 1182
Db 422 TGAGTGTGATGTAATCCATGAAAA 448

RESULT 13
AAS27206
ID AAS27206 standard; cdna; 430 BP.
XX
AC AAS27206;
XX
DT 07-NOV-2001 (first entry)
XX
cDNA encoding novel signal transduction pathway protein, Seq ID 241.
XX
XX Neuroprotective; cytostatic; dermatological; immunosuppressive; tumour;
XX antiinflammatory; anti-Hiv; antibacterial; antiinflammatory; cancer;
XX immune system disorder; rheumatoid arthritis; inflammatory condition;
XX organ transplant rejection; infection; hepatitis C; blood disorder;
XX sickle cell anaemia; hyperproliferative disorder; Gaucher's disease;
XX neurodegenerative disorder; Alzheimer's disease; Parkinson's disease;
XX chromosomal abnormality; Down syndrome; ischaemia; renal disorder;
XX cardiovascular; respiratory; wound healing; endocrine; Addison's disease;
XX reproductive system; gastrointestinal; liver disorder; AIDS; ss;
XX acquired immune deficiency syndrome.
XX
XX Homo sapiens.
XX
XX WO200154733-A1.
XX
XX 02-AUG-2001.
XX
XX 17-JAN-2001; 2001WO-US001132.
XX
XX 31-JAN-2000; 2000US-0179065P.
XX 04-FEB-2000; 2000US-0180628P.
XX 24-FEB-2000; 2000US-0184664P.
XX 02-MAR-2000; 2000US-0186350P.
XX 16-MAR-2000; 2000US-0189874P.
XX 17-MAR-2000; 2000US-0190076P.
XX 18-APR-2000; 2000US-0198123P.
XX 19-MAY-2000; 2000US-0205151P.
XX 07-JUN-2000; 2000US-0209467P.
XX 28-JUN-2000; 2000US-0214886P.
XX 30-JUN-2000; 2000US-0215135P.
XX 07-JUL-2000; 2000US-0216647P.
XX 07-JUL-2000; 2000US-0216880P.
XX 11-JUL-2000; 2000US-0217487P.
XX 11-JUL-2000; 2000US-0217496P.
XX 14-JUL-2000; 2000US-0218290P.
XX 26-JUL-2000; 2000US-0220963P.
XX 26-JUL-2000; 2000US-0220964P.
XX 14-AUG-2000; 2000US-0224518P.
XX 14-AUG-2000; 2000US-0224519P.
XX 14-AUG-2000; 2000US-0225213P.
PR 14-AUG-2000; 2000US-0225214P.
PR 14-AUG-2000; 2000US-0225266P.
PR 14-AUG-2000; 2000US-0225267P.
PR 14-AUG-2000; 2000US-0225268P.
PR 14-AUG-2000; 2000US-0225270P.
PR 14-AUG-2000; 2000US-0225447P.
PR 14-AUG-2000; 2000US-0225757P.
PR 14-AUG-2000; 2000US-0225758P.
PR 14-AUG-2000; 2000US-0225759P.
PR 14-AUG-2000; 2000US-0226279P.
PR 22-AUG-2000; 2000US-0226681P.
PR 22-AUG-2000; 2000US-0226682P.
PR 22-AUG-2000; 2000US-0227182P.
PR 23-AUG-2000; 2000US-0227009P.
PR 30-AUG-2000; 2000US-0228924P.
PR 01-SEP-2000; 2000US-0229287P.
PR 01-SEP-2000; 2000US-0229343P.
PR 01-SEP-2000; 2000US-0229344P.
PR 01-SEP-2000; 2000US-0229345P.
PR 05-SEP-2000; 2000US-0229509P.
PR 05-SEP-2000; 2000US-0229513P.
PR 06-SEP-2000; 2000US-0230437P.
PR 06-SEP-2000; 2000US-0230438P.
PR 08-SEP-2000; 2000US-0231242P.
PR 08-SEP-2000; 2000US-0231243P.
PR 08-SEP-2000; 2000US-0231244P.
PR 08-SEP-2000; 2000US-0231413P.
PR 08-SEP-2000; 2000US-0231414P.
PR 08-SEP-2000; 2000US-0232080P.
PR 08-SEP-2000; 2000US-0232081P.
PR 12-SEP-2000; 2000US-0231968P.
PR 14-SEP-2000; 2000US-0232397P.
PR 14-SEP-2000; 2000US-0232398P.
PR 14-SEP-2000; 2000US-0232399P.
PR 14-SEP-2000; 2000US-0232400P.
PR 14-SEP-2000; 2000US-0232401P.
PR 14-SEP-2000; 2000US-0233063P.
PR 14-SEP-2000; 2000US-0233064P.
PR 14-SEP-2000; 2000US-0233065P.
PR 21-SEP-2000; 2000US-0234223P.
PR 21-SEP-2000; 2000US-0234274P.
PR 25-SEP-2000; 2000US-0234997P.
PR 25-SEP-2000; 2000US-0234998P.
PR 26-SEP-2000; 2000US-0235484P.
PR 27-SEP-2000; 2000US-0235834P.
PR 27-SEP-2000; 2000US-0235836P.
PR 29-SEP-2000; 2000US-0236327P.
PR 29-SEP-2000; 2000US-0236367P.
PR 29-SEP-2000; 2000US-0236368P.
PR 29-SEP-2000; 2000US-0236369P.
PR 29-SEP-2000; 2000US-0236370P.
PR 02-OCT-2000; 2000US-0236802P.
PR 02-OCT-2000; 2000US-0237037P.
PR 02-OCT-2000; 2000US-0237038P.
PR 02-OCT-2000; 2000US-0237039P.
PR 13-OCT-2000; 2000US-0237040P.
PR 13-OCT-2000; 2000US-0239935P.
PR 13-OCT-2000; 2000US-0239937P.
PR 20-OCT-2000; 2000US-0240960P.
PR 20-OCT-2000; 2000US-0241221P.
PR 20-OCT-2000; 2000US-0241785P.
PR 20-OCT-2000; 2000US-0241786P.
PR 20-OCT-2000; 2000US-0241787P.
PR 20-OCT-2000; 2000US-0241808P.
PR 20-OCT-2000; 2000US-0241809P.
PR 20-OCT-2000; 2000US-0241826P.
PR 01-NOV-2000; 2000US-0244617P.
PR 08-NOV-2000; 2000US-0246474P.
PR 08-NOV-2000; 2000US-0246475P.
PR 08-NOV-2000; 2000US-0246476P.
PR 08-NOV-2000; 2000US-0246477P.
PR 08-NOV-2000; 2000US-0246478P.
PR 08-NOV-2000; 2000US-0246523P.
```

CC	B-cell responsiveness to pathogens, activators of T-cells, to induce
CC	higher affinity antibodies, and as a means to induce tumour proliferation
CC	in pathologies e.g. acquired immune deficiency syndrome (AIDS). AAS26976-
CC	AAS27850 represent novel signal transduction pathway protein coding
CC	sequences and PCR primers of the invention
XX	
Query Match	27.3%; Score 395.4; DB 4; Length 430;
Best Local Similarity	97.9%; Pred. No. 4.2e-105;
Matches 418; Conservative	1; Mismatches 6; Indels 2; Gaps 2;
QY	758 CTGAGAGATATCCAAAAGAACTAAATGCCATCATCGGAAGCATGTTGACAGGATCCTT 817
DB	1 CTGAGAGATATCCAAAAGAACTAAATGCCATCATCGGAAGCATGTTGACAGGATCCTT 60
QY	818 CATTAAAGACCATCTGCTATCGAAATTTTAAAAATCCCTTACCTTTGATGAGCAGCTACAGA 877
DB	61 CATTAAAGACCATCTGCTATCGAAATTTTAAAAATCCCTTACCTTTGATGAGCAGCTACAGA 120
QY	878 ACCTAATGTGTAGATATTTCAGAAATGACTCTGGAAAGACAAAATTTGGATTGTTCAGAGG 937
DB	121 ACCTAATGTGTAGATATTTCAGAAATGACTCTGGAAAGACAAAATTTGGATTGTTCAGAGG 180
QY	938 AGGCTGCTCATATATTAATGCTATGCAAAAAGATCCACTGCAGACTCTGAGGCGAC 997
DB	181 AGGCTGCTCATATATTAATGCTATGCAAAAAGATCCACTGCAGACTCTGAGGCGAC 240
QY	998 TGTCAAGAGTACAGAAAATGACGCCAAGAGAAAGATCGCGCTGAGGAAGTCCAGG-CG 1056
DB	241 TGTCAAGAGTACAGAAAATGACGCCAAGAGAAAGATCGCGCTGAGGAAGTCCAGGCGC 300
QY	1057 GCTGATGAGAAAGCCAGCAAGCTGAAAAGATTCTGGAGAAATATGAGAAATATAGC 1116
DB	301 GCTGATGAGAAAGCCAGCAAGCTGAAAAGATTCTGGAGAAATATGAGAAATATAGC 360
QY	1117 AAACGAATGCAAGAAATTCAGATCTCGGAACCTTTTCAGCAGCTGAGTGTGTG-ATGTA 1175
DB	361 AAACGAATGCAAGAAATTCAGATCTCGGAACCTTTTCAGCAGCTGAGTGTGTGTAATGTA 420
QY	1176 TGAAGAAA 1182
DB	421 TGAAGAAA 427
RESULT 14	
AAK8454	
ID	AAK8454 standard; cDNA; 430 BP.
XX	
AC	AAK8454;
XX	
DT	05-NOV-2001 (first entry)
XX	
DE	Human digestive system antigen coding sequence SEQ ID NO: 770.
XX	
KW	Human; digestive system antigen; gene therapy; cancer; appendicitis;
KW	ulcerative colitis; infection; Hirschsprung's disease; chronic colitis;
XX	digestive system disorder; Meckel's diverticulum; ss.
OS	Homo sapiens.
XX	
PN	WO200155314-A2.
XX	
PD	02-AUG-2001.
XX	
PF	17-JAN-2001; 2001WO-US001324.
XX	
PR	31-JAN-2000; 2000US-0179065P.
PR	04-FEB-2000; 2000US-0180628P.
PR	24-FEB-2000; 2000US-0184664P.
PR	02-MAR-2000; 2000US-0186350P.
PR	16-MAR-2000; 2000US-0189874P.
PR	17-MAR-2000; 2000US-0190076P.
PR	18-APR-2000; 2000US-0198123P.

PR	08-NOV-2000; 2000US-0246524P.
PR	08-NOV-2000; 2000US-0246525P.
PR	08-NOV-2000; 2000US-0246526P.
PR	08-NOV-2000; 2000US-0246527P.
PR	08-NOV-2000; 2000US-0246528P.
PR	08-NOV-2000; 2000US-0246529P.
PR	08-NOV-2000; 2000US-0246609P.
PR	08-NOV-2000; 2000US-0246610P.
PR	08-NOV-2000; 2000US-0246611P.
PR	08-NOV-2000; 2000US-0246613P.
PR	17-NOV-2000; 2000US-0249207P.
PR	17-NOV-2000; 2000US-0249208P.
PR	17-NOV-2000; 2000US-0249209P.
PR	17-NOV-2000; 2000US-0249210P.
PR	17-NOV-2000; 2000US-0249211P.
PR	17-NOV-2000; 2000US-0249212P.
PR	17-NOV-2000; 2000US-0249213P.
PR	17-NOV-2000; 2000US-0249214P.
PR	17-NOV-2000; 2000US-0249215P.
PR	17-NOV-2000; 2000US-0249216P.
PR	17-NOV-2000; 2000US-0249217P.
PR	17-NOV-2000; 2000US-0249218P.
PR	17-NOV-2000; 2000US-0249244P.
PR	17-NOV-2000; 2000US-0249245P.
PR	17-NOV-2000; 2000US-0249264P.
PR	17-NOV-2000; 2000US-0249265P.
PR	17-NOV-2000; 2000US-0249297P.
PR	17-NOV-2000; 2000US-0249300P.
PR	01-DEC-2000; 2000US-0250160P.
PR	01-DEC-2000; 2000US-0250391P.
PR	05-DEC-2000; 2000US-0251030P.
PR	05-DEC-2000; 2000US-0251988P.
PR	05-DEC-2000; 2000US-0256719P.
PR	06-DEC-2000; 2000US-0251479P.
PR	08-DEC-2000; 2000US-0251856P.
PR	08-DEC-2000; 2000US-0251868P.
PR	08-DEC-2000; 2000US-0251869P.
PR	08-DEC-2000; 2000US-0251989P.
PR	11-DEC-2000; 2000US-0251990P.
PR	05-JAN-2001; 2001US-0259678P.
PA	(HUMA-) HUMAN GENOME SCI INC.
PI	Rosen CA, Barash SC, Ruben SM;
XX	
DR	WPI; 2001-465460/50.
XX	
DE	P-PSDB; AAU17289.
XX	
PT	Novel polypeptides useful for diagnosing, treating, preventing and/or
PT	prognosing disorders related to the proteins, including cancers, immune
XX	disorders and neuronal disorders.
XX	
PS	Claim 1; SEQ ID NO 241; 880pp; English.
XX	
CC	The invention relates to novel isolated polypeptides (I), and
CC	polynucleotides (II). (I), (II) and the antibody to (I) are useful for
CC	diagnosing, preventing and treating diseases including immune system
CC	disorders (e.g. congenital and acquired immunodeficiencies, autoimmune
CC	disorders (e.g. rheumatoid arthritis), inflammatory conditions, organ
CC	transplant rejections and graft versus host disease, infectious diseases
CC	(e.g. hepatitis C), bleeding disorders, haemoglobin abnormalities and
CC	other blood-related disorders (sickle cell anaemia), myeloproliferative
CC	disorders, primary haematopoietic disorders, hyperproliferative disorders
CC	(e.g. Gaucher's disease and cancer), neurodegenerative disorders (e.g.
CC	Alzheimer's disease, Parkinson's disease), chromosomal abnormalities
CC	(Down syndrome), ischaemic injury (e.g. stroke), renal disorders (e.g.
CC	glomerulonephritis), cardiovascular disorders (e.g. arrhythmia),
CC	respiratory disorders, dermatological disorders, in wound healing,
CC	epithelial cell proliferation, endocrine disorders (e.g. Addison's
CC	disease), reproductive system disorders, gastrointestinal disorder
CC	(inflammatory disorders), liver disorders (cirrhosis), as stimulators of



PR 19-MAY-2000; 2000US-0205515P.  
PR 07-JUN-2000; 2000US-0209467P.  
PR 28-JUN-2000; 2000US-0214886P.  
PR 30-JUN-2000; 2000US-0215135P.  
PR 07-JUL-2000; 2000US-0216647P.  
PR 07-JUL-2000; 2000US-0216880P.  
PR 11-JUL-2000; 2000US-0217487P.  
PR 11-JUL-2000; 2000US-0217496P.  
PR 14-JUL-2000; 2000US-0218290P.  
PR 14-JUL-2000; 2000US-0220963P.  
PR 26-JUL-2000; 2000US-0220964P.  
PR 26-JUL-2000; 2000US-0220964P.  
PR 14-AUG-2000; 2000US-0224518P.  
PR 14-AUG-2000; 2000US-0224519P.  
PR 14-AUG-2000; 2000US-0225213P.  
PR 14-AUG-2000; 2000US-0225214P.  
PR 14-AUG-2000; 2000US-0225266P.  
PR 14-AUG-2000; 2000US-0225267P.  
PR 14-AUG-2000; 2000US-0225268P.  
PR 14-AUG-2000; 2000US-0225270P.  
PR 14-AUG-2000; 2000US-0225447P.  
PR 14-AUG-2000; 2000US-0225757P.  
PR 14-AUG-2000; 2000US-0225758P.  
PR 14-AUG-2000; 2000US-0225759P.  
PR 18-AUG-2000; 2000US-0226279P.  
PR 22-AUG-2000; 2000US-0226681P.  
PR 22-AUG-2000; 2000US-0226686P.  
PR 22-AUG-2000; 2000US-0227182P.  
PR 23-AUG-2000; 2000US-0227009P.  
PR 30-AUG-2000; 2000US-0228924P.  
PR 01-SEP-2000; 2000US-0229287P.  
PR 01-SEP-2000; 2000US-0229343P.  
PR 01-SEP-2000; 2000US-0229344P.  
PR 01-SEP-2000; 2000US-0229345P.  
PR 05-SEP-2000; 2000US-0229509P.  
PR 06-SEP-2000; 2000US-0229513P.  
PR 06-SEP-2000; 2000US-0230437P.  
PR 08-SEP-2000; 2000US-0230438P.  
PR 08-SEP-2000; 2000US-0231242P.  
PR 08-SEP-2000; 2000US-0231243P.  
PR 08-SEP-2000; 2000US-0231244P.  
PR 08-SEP-2000; 2000US-0231413P.  
PR 08-SEP-2000; 2000US-0231414P.  
PR 08-SEP-2000; 2000US-0232080P.  
PR 08-SEP-2000; 2000US-0232081P.  
PR 12-SEP-2000; 2000US-0231968P.  
PR 14-SEP-2000; 2000US-0232397P.  
PR 14-SEP-2000; 2000US-0232398P.  
PR 14-SEP-2000; 2000US-0232399P.  
PR 14-SEP-2000; 2000US-0232400P.  
PR 14-SEP-2000; 2000US-0233063P.  
PR 14-SEP-2000; 2000US-0233064P.  
PR 14-SEP-2000; 2000US-0233065P.  
PR 21-SEP-2000; 2000US-0234223P.  
PR 21-SEP-2000; 2000US-0234274P.  
PR 25-SEP-2000; 2000US-0234997P.  
PR 25-SEP-2000; 2000US-0234998P.  
PR 26-SEP-2000; 2000US-0235484P.  
PR 27-SEP-2000; 2000US-0235834P.  
PR 27-SEP-2000; 2000US-0235836P.  
PR 29-SEP-2000; 2000US-0236327P.  
PR 29-SEP-2000; 2000US-0236327P.  
PR 29-SEP-2000; 2000US-0236368P.  
PR 29-SEP-2000; 2000US-0236369P.  
PR 29-SEP-2000; 2000US-0236370P.  
PR 02-OCT-2000; 2000US-0236802P.  
PR 02-OCT-2000; 2000US-0237037P.  
PR 02-OCT-2000; 2000US-0237038P.  
PR 02-OCT-2000; 2000US-0237039P.  
PR 02-OCT-2000; 2000US-0237040P.  
PR 13-OCT-2000; 2000US-0239935P.  
PR 13-OCT-2000; 2000US-0239937P.  
PR 20-OCT-2000; 2000US-0240960P.

20-OCT-2000; 2000US-0241221P.  
PR 20-OCT-2000; 2000US-0241785P.  
PR 20-OCT-2000; 2000US-0241786P.  
PR 20-OCT-2000; 2000US-0241787P.  
PR 20-OCT-2000; 2000US-0241808P.  
PR 20-OCT-2000; 2000US-0241809P.  
PR 20-OCT-2000; 2000US-0241826P.  
PR 01-NOV-2000; 2000US-0244617P.  
PR 08-NOV-2000; 2000US-0246474P.  
PR 08-NOV-2000; 2000US-0246475P.  
PR 08-NOV-2000; 2000US-0246476P.  
PR 08-NOV-2000; 2000US-0246477P.  
PR 08-NOV-2000; 2000US-0246478P.  
PR 08-NOV-2000; 2000US-0246523P.  
PR 08-NOV-2000; 2000US-0246524P.  
PR 08-NOV-2000; 2000US-0246525P.  
PR 08-NOV-2000; 2000US-0246526P.  
PR 08-NOV-2000; 2000US-0246527P.  
PR 08-NOV-2000; 2000US-0246528P.  
PR 08-NOV-2000; 2000US-0246529P.  
PR 08-NOV-2000; 2000US-0246609P.  
PR 08-NOV-2000; 2000US-0246610P.  
PR 08-NOV-2000; 2000US-0246611P.  
PR 08-NOV-2000; 2000US-0246613P.  
PR 17-NOV-2000; 2000US-0249207P.  
PR 17-NOV-2000; 2000US-0249208P.  
PR 17-NOV-2000; 2000US-0249209P.  
PR 17-NOV-2000; 2000US-0249210P.  
PR 17-NOV-2000; 2000US-0249211P.  
PR 17-NOV-2000; 2000US-0249212P.  
PR 17-NOV-2000; 2000US-0249213P.  
PR 17-NOV-2000; 2000US-0249214P.  
PR 17-NOV-2000; 2000US-0249215P.  
PR 17-NOV-2000; 2000US-0249216P.  
PR 17-NOV-2000; 2000US-0249217P.  
PR 17-NOV-2000; 2000US-0249218P.  
PR 17-NOV-2000; 2000US-0249244P.  
PR 17-NOV-2000; 2000US-0249245P.  
PR 17-NOV-2000; 2000US-0249264P.  
PR 17-NOV-2000; 2000US-0249265P.  
PR 17-NOV-2000; 2000US-0249297P.  
PR 17-NOV-2000; 2000US-0249299P.  
PR 17-NOV-2000; 2000US-0249300P.  
PR 01-DEC-2000; 2000US-0250160P.  
PR 01-DEC-2000; 2000US-0250391P.  
PR 05-DEC-2000; 2000US-0251030P.  
PR 05-DEC-2000; 2000US-0251988P.  
PR 05-DEC-2000; 2000US-0256719P.  
PR 06-DEC-2000; 2000US-0251479P.  
PR 08-DEC-2000; 2000US-0251856P.  
PR 08-DEC-2000; 2000US-0251868P.  
PR 08-DEC-2000; 2000US-0251869P.  
PR 08-DEC-2000; 2000US-0251869P.  
PR 08-DEC-2000; 2000US-0251989P.  
PR 08-DEC-2000; 2000US-0251990P.  
PR 11-DEC-2000; 2000US-0254097P.  
PR 05-JAN-2001; 2001US-0259678P.

(HUMA-) HUMAN GENOME SCI INC.

Rosen CA, Barash SC, Ruben SM;

WPI; 2001-502630/55.

P-PSDB; AAM92681.

Polynucleotides encoding digestive system antigens, useful for diagnosing, treating, preventing and/or prognosing disorders of the digestive system, particularly cancer and cancer metastases.

Claim 1; SEQ ID NO 770; 986pp; English.

The present invention provides the protein and coding sequences of a number of human digestive system antigens. These can be used in the diagnosis, treatment and prevention of digestive system disorders,

CC including cancer, Meckel's diverticulum, bacterial or parasitic  
 CC infections, appendicitis, Hirschsprung's disease, chronic colitis or  
 CC ulcerative colitis. The present sequence is a cDNA encoding a digestive  
 CC system antigen of the invention  
 XX  
 SQ Sequence 430 BP; 161 A; 76 C; 98 G; 90 T; 0 U; 5 Other;

Query Match 27.3%; Score 395.4; DB 4; Length 430;  
 Best Local Similarity 97.9%; Pred. No. 4.2e-105;  
 Matches 418; Conservative 1; Mismatches 6; Indels 2; Gaps 2;

Qy 758 CTGAGAGATATCCAAAGAACTAAATGCCATCATGGAAGCATGTTGAAACAAGATCCTT 817  
 Db 1 CTGAGAGATATCCAAAGAACTAAATGCCATCATGGAAGCATGTTGAAACAAGATCCTT 60

Qy 818 CATTAGACCATCTGCTATCGAATTTTAAATCCCTTACCTTGATGACGACTACAGA 877  
 Db 61 CATTAGACCATCTGCTATCGAATTTTAAATCCCTTACCTTGATGACGACTACAGA 120

Qy 878 ACCTAATGTGTAGATATTTCAGAAATGACTCTGGAAGACAAAATTTGGATTGTGAGAAGG 937  
 Db 121 ACCTAATGTGTAGATATTTCAGAAATGACTCTGGAAGACAAAATTTGGATTGTGAGAAGG 180

Qy 938 AGGCTGCTCATATTAATTAATGCGATCGAAGAAAGATCCACCTGCAGACTCTGAGGGCAC 997  
 Db 181 AGGCTGCTCATATTAATTAATGCGATCGAAGAAAGATCCACCTGCAGACTCTGAGGGCAC 240

Qy 998 TGTGAGAAGTACAGAAATACGCAAGAGAAAGATCGCGCTGAGAACTCCAGG-CG 1056  
 Db 241 TGTGAGAAGTACAGAAATACGCAAGAGAAAGATCGCGCTGAGAACTCCAGGCGC 300

Qy 1057 GCTGATGAGAAGCCAGGAAGCTGAAAAGATTTGGAAGAAAATATGAAGAAATAGC 1116  
 Db 301 GCTGATGAGAAGCCAGGAAGCTGAAAAGATTTGGAAGAAAATATGAAGAAATAGC 360

Qy 1117 AAACGAATGCAAGATTTGAGATCTCGGAATCTTCAGCAGCTGAGTGTG-ATGTACTCCA 1175  
 Db 361 AAACGAATGCAAGATTTGAGATCTCGGAATCTTCAGCAGCTGAGTGTGATGTACTTCA 420

Qy 1176 TGAATAA 1182  
 Db 421 TGAATAA 427

RESULT 15  
 ABK43786  
 ID ABK43786 standard; cDNA; 430 BP.  
 XX  
 AC ABK43786;  
 XX  
 DT 05-JUN-2002 (first entry)  
 XX  
 DE DNA encoding novel central nervous system protein #366.  
 XX  
 KW Central nervous system; CNS; autoimmune disease; rheumatoid arthritis;  
 KW hyperproliferative disorder; neoplasm; cardiovascular disorder;  
 KW cardiac arrest; cerebrovascular disorder; ischemia; angiogenesis;  
 KW nervous system disorder; Alzheimer's disease; AIDS; ocular disorder;  
 KW acquired immunodeficiency virus; dysphagia; gastrointestinal disorder;  
 KW adenocarcinoma; reproductive system disorder; testicular feminisation;  
 KW endocrine disorder; diabetes; cancer; leukaemia; neovascularisation;  
 KW respiratory disorder; renal disorder; kidney failure; blood disorder;  
 KW myocardial infarction; wound healing; cell proliferation; skin aging;  
 KW food additive; food preservative; gene therapy; gene; ss.  
 OS Homo sapiens.  
 XX  
 PN WO200155318-A2.  
 XX  
 PD 02-AUG-2001.  
 XX  
 PF 17-JAN-2001; 2001WO-US001332.  
 XX

PR 31-JAN-2000; 2000US-0179065P.  
 PR 04-FEB-2000; 2000US-0180628P.  
 PR 24-FEB-2000; 2000US-0184644P.  
 PR 02-MAR-2000; 2000US-0186350P.  
 PR 16-MAR-2000; 2000US-0189874P.  
 PR 17-MAR-2000; 2000US-0190076P.  
 PR 18-APR-2000; 2000US-0198123P.  
 PR 19-MAY-2000; 2000US-0205515P.  
 PR 07-JUN-2000; 2000US-0209467P.  
 PR 28-JUN-2000; 2000US-0214886P.  
 PR 30-JUN-2000; 2000US-0215135P.  
 PR 07-JUL-2000; 2000US-0216647P.  
 PR 07-JUL-2000; 2000US-0216880P.  
 PR 11-JUL-2000; 2000US-0217487P.  
 PR 11-JUL-2000; 2000US-0217496P.  
 PR 14-JUL-2000; 2000US-0218290P.  
 PR 26-JUL-2000; 2000US-0220963P.  
 PR 26-JUL-2000; 2000US-0220964P.  
 PR 14-AUG-2000; 2000US-0224518P.  
 PR 14-AUG-2000; 2000US-0224519P.  
 PR 14-AUG-2000; 2000US-0225213P.  
 PR 14-AUG-2000; 2000US-0225214P.  
 PR 14-AUG-2000; 2000US-0225266P.  
 PR 14-AUG-2000; 2000US-0225267P.  
 PR 14-AUG-2000; 2000US-0225268P.  
 PR 14-AUG-2000; 2000US-0225270P.  
 PR 14-AUG-2000; 2000US-0225447P.  
 PR 14-AUG-2000; 2000US-0225757P.  
 PR 14-AUG-2000; 2000US-0225758P.  
 PR 14-AUG-2000; 2000US-0225759P.  
 PR 18-AUG-2000; 2000US-0226279P.  
 PR 22-AUG-2000; 2000US-0226681P.  
 PR 22-AUG-2000; 2000US-0226888P.  
 PR 23-AUG-2000; 2000US-0227182P.  
 PR 30-AUG-2000; 2000US-0228924P.  
 PR 01-SEP-2000; 2000US-0229287P.  
 PR 01-SEP-2000; 2000US-0229343P.  
 PR 01-SEP-2000; 2000US-0229344P.  
 PR 01-SEP-2000; 2000US-0229345P.  
 PR 05-SEP-2000; 2000US-0229509P.  
 PR 05-SEP-2000; 2000US-0229513P.  
 PR 06-SEP-2000; 2000US-0230437P.  
 PR 08-SEP-2000; 2000US-0230438P.  
 PR 08-SEP-2000; 2000US-0231242P.  
 PR 08-SEP-2000; 2000US-0231243P.  
 PR 08-SEP-2000; 2000US-0231244P.  
 PR 08-SEP-2000; 2000US-0231413P.  
 PR 08-SEP-2000; 2000US-0231414P.  
 PR 08-SEP-2000; 2000US-0232080P.  
 PR 12-SEP-2000; 2000US-0232081P.  
 PR 14-SEP-2000; 2000US-0231968P.  
 PR 14-SEP-2000; 2000US-0232397P.  
 PR 14-SEP-2000; 2000US-0232398P.  
 PR 14-SEP-2000; 2000US-0232399P.  
 PR 14-SEP-2000; 2000US-0232400P.  
 PR 14-SEP-2000; 2000US-0232401P.  
 PR 14-SEP-2000; 2000US-0233063P.  
 PR 14-SEP-2000; 2000US-0233064P.  
 PR 14-SEP-2000; 2000US-0233065P.  
 PR 21-SEP-2000; 2000US-0234223P.  
 PR 25-SEP-2000; 2000US-0234274P.  
 PR 25-SEP-2000; 2000US-0234997P.  
 PR 26-SEP-2000; 2000US-0234988P.  
 PR 26-SEP-2000; 2000US-0235484P.  
 PR 27-SEP-2000; 2000US-0235834P.  
 PR 27-SEP-2000; 2000US-0235836P.  
 PR 29-SEP-2000; 2000US-0236367P.  
 PR 29-SEP-2000; 2000US-0236368P.  
 PR 29-SEP-2000; 2000US-0236369P.  
 PR 29-SEP-2000; 2000US-0236370P.  
 PR 02-OCT-2000; 2000US-0236802P.

PR 02-OCT-2000; 2000US-0217037P.  
PR 02-OCT-2000; 2000US-0237038P.  
PR 02-OCT-2000; 2000US-0237039P.  
PR 02-OCT-2000; 2000US-0237040P.  
PR 13-OCT-2000; 2000US-0239935P.  
PR 13-OCT-2000; 2000US-0239937P.  
PR 20-OCT-2000; 2000US-0240960P.  
PR 20-OCT-2000; 2000US-0241221P.  
PR 20-OCT-2000; 2000US-0241785P.  
PR 20-OCT-2000; 2000US-0241786P.  
PR 20-OCT-2000; 2000US-0241787P.  
PR 20-OCT-2000; 2000US-0241808P.  
PR 20-OCT-2000; 2000US-0241809P.  
PR 20-OCT-2000; 2000US-0241826P.  
PR 01-NOV-2000; 2000US-0244617P.  
PR 08-NOV-2000; 2000US-0246474P.  
PR 08-NOV-2000; 2000US-0246475P.  
PR 08-NOV-2000; 2000US-0246476P.  
PR 08-NOV-2000; 2000US-0246477P.  
PR 08-NOV-2000; 2000US-0246478P.  
PR 08-NOV-2000; 2000US-0246523P.  
PR 08-NOV-2000; 2000US-0246524P.  
PR 08-NOV-2000; 2000US-0246525P.  
PR 08-NOV-2000; 2000US-0246526P.  
PR 08-NOV-2000; 2000US-0246527P.  
PR 08-NOV-2000; 2000US-0246528P.  
PR 08-NOV-2000; 2000US-0246529P.  
PR 08-NOV-2000; 2000US-0246609P.  
PR 08-NOV-2000; 2000US-0246610P.  
PR 08-NOV-2000; 2000US-0246611P.  
PR 08-NOV-2000; 2000US-0246613P.  
PR 17-NOV-2000; 2000US-0249207P.  
PR 17-NOV-2000; 2000US-0249208P.  
PR 17-NOV-2000; 2000US-0249209P.  
PR 17-NOV-2000; 2000US-0249210P.  
PR 17-NOV-2000; 2000US-0249211P.  
PR 17-NOV-2000; 2000US-0249212P.  
PR 17-NOV-2000; 2000US-0249213P.  
PR 17-NOV-2000; 2000US-0249214P.  
PR 17-NOV-2000; 2000US-0249215P.  
PR 17-NOV-2000; 2000US-0249216P.  
PR 17-NOV-2000; 2000US-0249217P.  
PR 17-NOV-2000; 2000US-0249218P.  
PR 17-NOV-2000; 2000US-0249219P.  
PR 17-NOV-2000; 2000US-0249244P.  
PR 17-NOV-2000; 2000US-0249245P.  
PR 17-NOV-2000; 2000US-0249246P.  
PR 17-NOV-2000; 2000US-0249255P.  
PR 17-NOV-2000; 2000US-0249257P.  
PR 17-NOV-2000; 2000US-0249299P.  
PR 17-NOV-2000; 2000US-0249300P.  
PR 01-DEC-2000; 2000US-0250160P.  
PR 01-DEC-2000; 2000US-0250391P.  
PR 05-DEC-2000; 2000US-0251030P.  
PR 05-DEC-2000; 2000US-0251988P.  
PR 06-DEC-2000; 2000US-0251795P.  
PR 08-DEC-2000; 2000US-0251856P.  
PR 08-DEC-2000; 2000US-0251868P.  
PR 08-DEC-2000; 2000US-0251869P.  
PR 08-DEC-2000; 2000US-0251899P.  
PR 08-DEC-2000; 2000US-0251990P.  
PR 11-DEC-2000; 2000US-0254097P.  
PR 05-JAN-2001; 2001US-0259678P.  
XX

(HUMA-) HUMAN GENOME SCI INC.

Rosen CA, Barash SC, Ruben SM;

WFI; 2001-581633/65.

P-PSDS; AAU87456.

PT New isolated nucleic acid encoding a protein for diagnosing, preventing, treating or ameliorating medical conditions and used as food additives or

PT Preservatives.  
XX Claim 1; SEQ ID NO 376; 837pp; English.  
XX  
CC The invention describes an isolated nucleic acid molecule (I) encoding a novel central nervous system protein. (II) and polypeptides (III) encoded by (I), are used to treat a medical condition and in diagnosis of a pathological condition. Disorders which are diagnosed or treated include autoimmune diseases e.g. rheumatoid arthritis, hyperproliferative disorders e.g. neoplasms of the breast or liver, cardiovascular disorders e.g. cardiac arrest, cerebrovascular disorders e.g. cerebral ischaemia, angiogenesis, nervous system disorders e.g. Alzheimer's disease and amyotrophic lateral sclerosis, infections caused by bacteria, viruses e.g. Acquired immunodeficiency virus (AIDS) and fungi, ocular disorders e.g. corneal infection, gastrointestinal disorders e.g. dysphagia, adenocarcinomas and irritable bowel syndrome, reproductive system disorders e.g. testicular feminisation, endocrine disorders e.g. diabetes and pituitary dwarfism, cancers and disorders at the cellular level e.g. leukaemia, disorders involving neovascularisation e.g. malignancies, respiratory disorders e.g. nonallergic rhinitis, renal disorders e.g. acute kidney failure and blood related disorders e.g. myocardial infarction. The polypeptides can also be used to aid wound healing and epithelial cell proliferation, to prevent skin aging due to sunburn, to maintain organs before transplant, for supporting cell culture of primary tissues, to regenerate tissues and in chemotaxis. The polypeptides can also be used as a food additive or preservative to increase or decrease storage capabilities, fat content, lipid, protein,

Query Match 27.3%; Score 395.4; DB 4; Length 430;

Best Local Similarity 97.9%; Pred. No. 4.2e-105;

Matches 418; Conservative 1; Mismatches 6; Indels 2; Gaps 2;

QY 758 CTGAGAGATATCCAAAGAACTAAATGCCATCATCGGAAGCATGTTGAACAGATCCTT 817  
Db 1 CTGAGAGATATCCAAAGAACTAAATGCCATCATCGGAAGCATGTTGAACAGATCCTT 60  
QY 818 CATTAGACCATCTGCTATCGAAATTTAAAAATCCCTTACCTTCATGAGCAGCTACAGA 877  
Db 61 CATTAGACCATCTGCTATCGAAATTTAAAAATCCCTTACCTTCATGAGCAGCTACAGA 120  
QY 878 ACCTAATGTGTAGATATTCAGAAATGACTCTGGAAGACAAATAATTTGGATTGTCAGAAG 937  
Db 121 ACCTAATGTGTAGATATTCAGAAATGACTCTGGAAGACAAATAATTTGGATTGTCAGAAG 180  
QY 938 AGGCTGCTCATATAATTAATGCGATGCAAAAAGGATCCACTGCAGACTCTGAGGGCAC 997  
Db 181 AGGCTGCTCATATAATTAATGCGATGCAAAAAGGATCCACTGCAGACTCTGAGGGCAC 240  
QY 998 TGTCAAGATACAGAAAATGAGCCCAAGAGAAAGATGCGGCTGAGAGCTCCAGG-CG 1056  
Db 241 TGTCAAGATACAGAAAATGAGCCCAAGAGAAAGATGCGGCTGAGAGCTCCAGGGCG 300  
QY 1057 GCTGATGAGAAAGCCAGGAGCTGAAAAGATTGTTGGAAGAAAATATGAAGAAATAGC 1116  
Db 301 GCTGATGAGAAAGCCAGGAGCTGAAAAGATTGTTGGAAGAAAATATGAAGAAATAGC 360  
QY 1117 AAACGAATGCAAGAAATTCAGATCTCGGAATTTTCAGCAGCTGAGTGTG-ATGTACTCCA 1175  
Db 361 AAACGAATGCAAGAAATTCAGATCTCGGAATTTTCAGCAGCTGAGTGTGATGTTACTTCA 420  
QY 1176 TGAATAA 1182  
Db 421 TGAATAA 427

Search completed: October 2, 2004, 22:12:24

Job time : 623 secs

***This Page Blank (uspto)***